Title	Statistical Analysis of Dose-Response Relationship
Author(s)	山本, 英二
Citation	大阪大学, 1983, 博士論文
Version Type	VoR
URL	https://hdl.handle.net/11094/464
rights	
Note	

Statistical Analysis

of

Dose-Response Relationship

1983

Eiji Yamamoto

Okayama University of Science

Summary

Three topics in statistical analysis of dose-response relationship
are discussed. The first topic is the estimation method of a safe dose
in dose-response relationship. The hockey stick regression method and
the probit analysis are compared. Statistical analysis and computational
results suggest that the probit analysis is preferable. The second topic
is the conversion method of a safe dose which may be used for the long-
range environmental standard of NO_2 in Japan. Two conversion methods,
current and the author's, are compared. Under the assumption of the
lognormal model for NO_2 concentrations, the author's methd is shown by
analytical and computational results to be preferable. The third topic
is the statistical inference for the dose-response relationship in
stratified populations. A set of three X^2 tests is proposed for the
estimation and test of the linear regression line of adjusted prevalence
rates on study factor intensities.

Contents

1 Introduction

Along with the arrival of the big industrial society of mass production
and mass consumption, harmful influence of various and too much industrial
chemical compounds to the human beings as well as to the biological
environment has constituted a serious social problem. Many scientists
and engineers are engaged in researching and preventing such influence.
Statistical analysis of the dose-response relationship which reveals
statistically the mechanism of the harmful influence is one of most
important themes of biostatisticians. Many epidemiological surveys and
their analyses have been carried out. Such analyses are expected to
progress further both in theory and application. This paper takes up
three topics which have aroused wide social interest in Japan.

In Chapter 2, a safe dose in dose-response relationship is discussed.
It is concerned with administrative control over air pollutions, food
additives or feed additives. Many techniques and methods to estimate
the safe dose have been proposed and discussed. A convenient estimation
method is the hockey stick regression method using segmented lines.
The method described in Section 2.1 [1] is based on the assumption that

there exists a threshold value which defines a safe dose. The validity

of the assumption is difficult to be shown. An alternative method which

is not based on the assumption is the probit analysis described in Section

2.2 [3]. It introduces a risk level which defines a safe dose. Discussion

on the two methods is presented in Section 2.3.

Chapter 3 deals with the conversion of a safe dose [2]. It gives a

critical review of the current long-range environmental standard of NO_2,

which is stated in terms of the 98 percent value of day means in a year.

The 98 percent value is a converted value from the year mean which had

been adopted previously as a safe dose from the standpoint of medical

science. The conversion was done by using the regression line of 98

percent values of day means on year means. Under the assumption of the

lognormal model for NO_2 concentrations, we show that the current conversion

method is unacceptable and shold be replaced by a conversion method using

the linear regression model with the error term whose standard deviation

is proportional to the independent variable. The former is discussed in

Section 3.1 and the latter in Section 3.2. In our method, a tolerance

limit shold be recommended as a converted value. In Section 3.3, the

deviation from the lognormal model is evaluated and our model proves to

be robust. Throughout this chapter, the data on NO_2 concentrations at

Okayama in 1977 are used as an illustration.

In Chapter 4, statistical inference for the dose-response relationship

in stratified populations is taken up [4]. The adjusted prevalence rate

which is used by many biostatisticians proves to be a useful tool in our

study. The dose-response relationship between adjusted prevalence rates

and study factor intensities may be represented locally and approximately by a linear regression model. A set of three X^2 tests is proposed for an estimation and test procedure of the linear regression line in Section 4.1. It is shown that the same statistical inference as in the single stratum population is valid by introducing the adjusted prevalence rates and the adjusted sample sizes in stratified populations. In Section 4.2, the procedure is compared with usual statistical procedures. It is shown that three X^2 tests of the procedure are usually more powerful than the X^2 tests performed by summing up X^2 statistics in each stratum. In Section 4.3, the procedure is applied to the data involving the average concentrations of NO_2 and the prevalence rates of persistent cough and phlegm. Section 4.4 deals with an extension to the probit model and the logistic model.

REFERENCES

[1] 山本英二、竹島克郎、柳本武美 (*1978*)。 安全基準としての閾値とホッケー
ステイック回帰法、統計数理研究所彙報、第*25*巻、第*1*号、*29－40*。

[2] 山本英二 (*1980*)。 対数正規分布に基づく閾値の換算法とNO_2環境基準、
応用統計学、第*9*巻、第*3*号、*169－179*。

[3] Yanagimoto, T. and Yamamoto, E. (1979). Estimation of Safe Doses: Critical Review of the Hockey Stick Regression Method, *Environmental Health Perspectives,* Vol. 32, pp. 193-199.

[4] Yamamoto, E. (1983). Chi-Square Tests for Linear Trend in Adjusted Prevalence Rates, *J. Japan Statist. Soc.,* Vol. 13.

2 A Safe Dose in Dose-Response Relationship

Many techniques and methods to estimate a safe dose in dose-response relationship have been proposed and discussed. Estimation of the safe dose concerning various chemical compounds is important, though it is very difficult.

The hockey stick (HS) regression method is an interesting method proposed by Hasselblad et al. [10] to obtain the maximum no-adverse-health -effect concentration of photochemical oxidants. It is a kind of regression method using segmented curves [7], [8] and has attracted attention of many researchers. The method is based on the assumption that there exists a threshold value which defines a safe dose.

In this chapter, we study some properties of the HS regression method, especially its validity. For this purpose, the HS model is compared with other regression models such as the probit model. For the latter, a risk level is used to define a safe dose. A risk level was used by Mantel and Bryan [5] and has been supported by subsequent researchers. A safe dose does not mean here a dose which causes no harmful effects, and therefore should not be used as a standard by the administration as it stands. It should be considered rather as a criterion to estimate

the safety.

The data on which the current standard of SO_2 in Japan is partly based are analyzed and discussed. A reanalysis of relationship between photochemical oxidant and eye discomfort is given. Using two examples, we compare the HS model and the probit model. Some conclusions and suggestions are presented.

2.1 A THRESHOLD VALUE AND THE HOCKEY STICK REGRESSION METHOD

2.1.1 The HS Regression Method

Let $p(x)$ denote the population prevalence rate at a dose level x. The HS regression function as a dose-response curve is defined as follows. For some x_o,

$$(2.1) \qquad p(x) = \beta_0 \qquad for \quad x \leq x_o,$$
$$= \beta_1 + \beta_2 x \quad for \quad x > x_o.$$

This means that for a suitable dose x_o, $p(x)$ remains constant for any x less than x_o and increases linearly as x increases from x_o. The dose x_o is considered as a physiological threshold value, whereas β_0 represents a spontaneous or baseline response which is caused by background stimuli.

Suppose that the survey is done at N points and let x_i and y_i denote the dose level and the sample prevalence rate at the ith point. The HS model is that

(2.2) $$y_i = p(x_i) + e_i \qquad (i = 1, \ldots, N),$$

$$e_i \sim N(0, \sigma^2).$$

The HS regression method is the least square estimation method of x_o under the HS regression model. Sometimes a flat line β_0 and a linear line $\beta_1 + \beta_2 x$ are estimated by separated data. Data to estimate the former and the latter are considered as those of non-polluted and polluted areas, respectively. Generally, both lines are estimated simultaneously under the constraint that they are connected at the point x_o. Using the estimators $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$, the estimator $\hat{x}_o$ of x_o is defined by

(2.3) $$\hat{\beta}_0 = \hat{\beta}_1 + \hat{\beta}_2 \hat{x}_o.$$

An assumtion of the existence of a threshold value is necessary to consider x_o as a safe dose, on which the HS regression model is based. This assumption seems to be serious, since we have no proof of the existence of such a value for substances such as food additives and environmental pollutants which many human beings are exposed to. Generally the HS model is only an operational one to obtain some value as the safe dose.

The HS method is accepted by Japanese research workers, epidemiologists, who are interested in finding a relationship between concentration of air pollutants and prevalence rate of some disease from epidemiological surveys. Such a relationship is needed to obtain a criterion on which

air quality standard shold be based. An example is the current air

quality standard of SO_2 in Japan. The surveys were conducted in Osaka,

Akoh, and Yokkaichi cities [2]. A prevalence rate of positive simple

chronic bronchitis for each area was obtained. Questionnaires given

by British Medical Research Council [6] were made on respiratory symptoms.

Thus chronic bronchitis is defined as persistent cough and phlegm. The

prevalence rates and average concentrations of SO_2 during three years

are listed in Table 2.1.

Table 2.1   Average concentrations of SO_2 and
            prevalence rates of chronic bronchitis.

SO_2 mg/day/100 cm^2	Prevalence rate
0.21	0.035
0.28	0.033
0.27	0.031
0.15	0.030
0.15	0.029
0.14	0.027
0.13	0.027
0.14	0.025
3.4	0.078
2.75	0.059
2.75	0.052
2.1	0.048
1.6	0.038
1.55	0.037
1.15	0.032
1.0	0.027
0.9	0.024

Since the sample size in each area is not mentioned, we regard it as 2000 when necessary. The original analysis [2] is as follows. The survey areas are divided into two groups: eight areas are considered as non-polluted and nine areas as polluted. Let y_i denote the prevalence rate, and x_i the average concentration of SO_2 in the ith area. The HS regression method is used: β_0 is estimated by the data from the non-polluted areas, while β_1 and β_2 by the data from the polluted areas. The fitted regression line is given by

$$(2.4) \qquad \hat{\beta}_0 = 0.02963,$$

$$\hat{\beta}_1 + \hat{\beta}_2 x = 0.00765 + 0.01898x,$$

which is described in Fig. 2.1. From (2.3), $\hat{x}_o$ is estimated at 1.160.
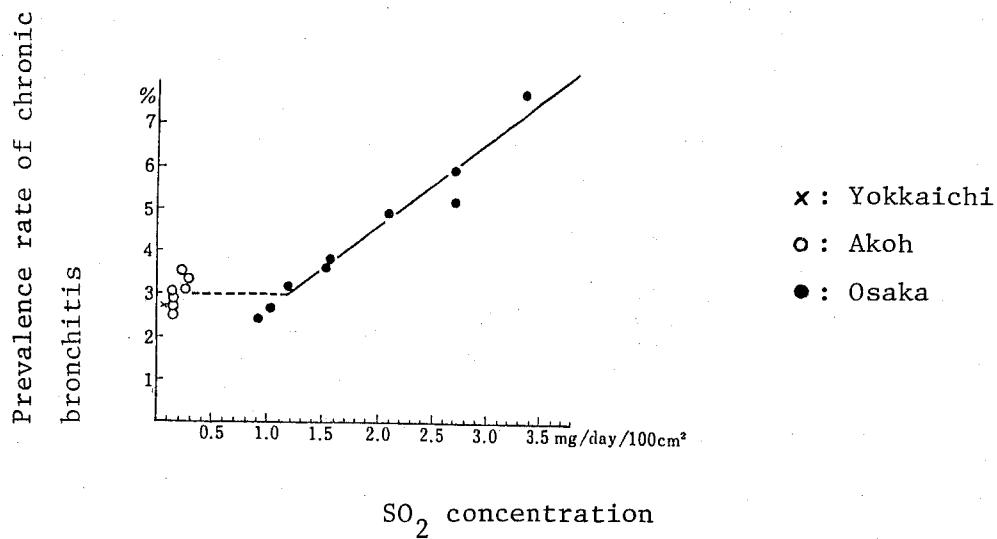
Fig. 2.1 Data of SO_2 concentrations and prevalence rates of chronic bronchitis.

The HS methods was studied originally in [10] and used to get a relationship between daily maximum hourly oxidant levels and daily symptom rates reported by student nurses in Los Angeles [11]. The summary data are cited in Table 2.2.

Table 2.2 Relationship of average daily percent of adjusted eye discomfort to photochemical oxidant levels.

Daily maximum hourly oxidant level (ppm)	No. of days	Average No. of nurses reporting daily	Average daily percent of adjusted eye discomfort (%)
≤0.04	229	64	5.0
0.05-0.08	184	59	5.4
0.09	35	58	5.6
0.10-0.14	176	62	5.9
0.15-0.19	144	58	6.9
0.20-0.24	63	60	9.2
0.25-0.29	25	60	11.2
0.30-0.39	9	67	17.8
0.40-0.50	3	53	31.8

Eye discomfort is a typical symptom caused by photochemical oxidants. The daily maximum hourly oxidant level is obtained as the midpoint of each interval shown in Table 2.2. The original authors used the HS regression method, estimating the parameters β_0, β_1, β_2 by the constrained least-square method:

(2.5)
$$\hat{\beta}_0 = 0.0541,$$

$$\hat{\beta}_1 + \hat{\beta}_2 x = -0.0172 + 0.491x.$$

The estimated threshold value is $\hat{x}_o = 0.145$. It is described in Fig. 2.2.

Average daily maximum hourly
photochemical oxidant level
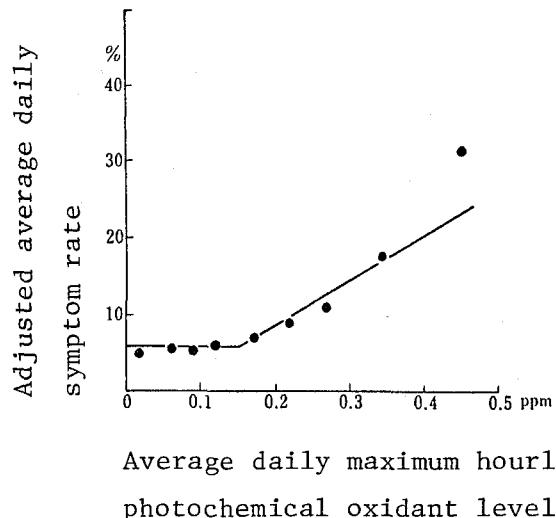
Fig. 2.2 Data of photochemical oxidant and eye discomfort

Table 2.1 is the case in which β_0 and (β_1, β_2) are estimated separately.
Table 2.2 is the case in which $(\beta_0, \beta_1, \beta_2)$ are estimated simultaneously.

2.1.2 Estimation of the Threshold Value

Under the HS model, we shall discuss two estimation methods of the
threshold value x_o. One is the case of separated data and the other
is the general case.

(1) Case of separated data

Data are separated into non-polluted areas and polluted areas, from
which β_0 and (β_1, β_2) are estimated separately. A defect of the original
analysis of Table 2.1 is its giving no evaluation for the variance of $\hat{x}_o$.
In order to consider $\hat{x}_o$ as a safe dose, the lower confidence limit of x_o
is preferable to the estimator itself.

We shall discuss the distribution of the estimator $\hat{x}_o$. Let (x_1, y_1),

..., (x_{N_1}, y_{N_1}), (x_{N_1+1}, y_{N_1+1}), ..., $(x_{N_1+N_2}, y_{N_1+N_2})$, $N = N_1 + N_2$, be

the given data. The former part of size N_1 is for a flat line β_0 and the latter part of size N_2 is for a linear line $\beta_1 + \beta_2 x$. In Table 2.1, $N_1 = 8$ and $N_2 = 9$. The HS model for the data is as follows.

$$(2.6) \qquad y = X_H(\beta_0, \beta_1, \beta_2)' + e,$$

where
$$X_H = \begin{pmatrix} 1 & 0 & 0 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1 & 0 & 0 \\ 0 & 1 & x_{N_1+1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 0 & 1 & x_{N_1+N_2} \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ y_{N_1} \\ y_{N_1+1} \\ \cdot \\ \cdot \\ y_{N_1+N_2} \end{pmatrix}, \quad e = \begin{pmatrix} e_1 \\ \cdot \\ \cdot \\ e_{N_1} \\ e_{N_1+1} \\ \cdot \\ \cdot \\ e_{N_1+N_2} \end{pmatrix}.$$

Then, it follows from the least square method that

$$(2.7) \qquad \hat{x}_o = (\bar{y}_1 - \bar{y}_2)\frac{\displaystyle\sum_{j=1}^{N_2}(x_{N_1+j} - \bar{x}_2)^2}{\displaystyle\sum_{j=1}^{N_2}(x_{N_1+j} - \bar{x}_2)(y_{N_1+j} - \bar{y}_2)} + \bar{x}_2,$$

where $\bar{y}_1 = \dfrac{1}{N_1}\displaystyle\sum_{i=1}^{N_1} y_i$, $\bar{y}_2 = \dfrac{1}{N_2}\displaystyle\sum_{j=1}^{N_2} y_{N_1+j}$ and $\bar{x}_2 = \dfrac{1}{N_2}\displaystyle\sum_{j=1}^{N_2} x_{N_1+j}$.

We examine the distribution of $\hat{x}_o$ in terms of the parameters $(\beta_0, \beta_1, \beta_2, \sigma^2)$. The linear regression theory gives

$$(2.8) \qquad \begin{pmatrix} \hat{\beta}_0 - \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \beta_0 - \beta_1 \\ \beta_2 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}(X_H'X_H)^{-1}\begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}' \right).$$

Then, the distribution of $\hat{x}_o = (\hat{\beta}_0 - \hat{\beta}_1)/\hat{\beta}_2$ is not normal, but a

noncentral-noncentral t distribution of one degree of freedom. It has

no mean nor variance. The following explains the above contention.

Assume that

$$(2.9) \qquad \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ & \sigma_2^2 \end{pmatrix} \right),$$

then we have

$$(2.10) \qquad W = \frac{X_1}{X_2} \sim \frac{\sigma_1\rho}{\sigma_2} + \frac{(\mu_1 - \frac{\sigma_1}{\sigma_2}\rho\mu_2) + \sigma_1\sqrt{1 - \rho^2}\, Z_1}{\mu_2 + \sigma_2 Z_2},$$

where random variables Z_1 and Z_2 are independently and identically distributed

in $N(0, 1)$. Fieller [4] gave the exact expression of the probability density

function of W. He also gave approximate expressions of the probability

density function and the distribution function of W when the coefficient

of variation of X_2 is small. They are given in Fig. 2.3 and Fig. 2.4,

using the estimates $(\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\sigma}^2)$ from Table 2.1 instead of true values.

Because the parameters are not known, we do not give the true distribution

of $\hat{x}_o$, but give a general view only. For the numerical computation of

the distribution function, the L-function in statistical tables [3] was

used.
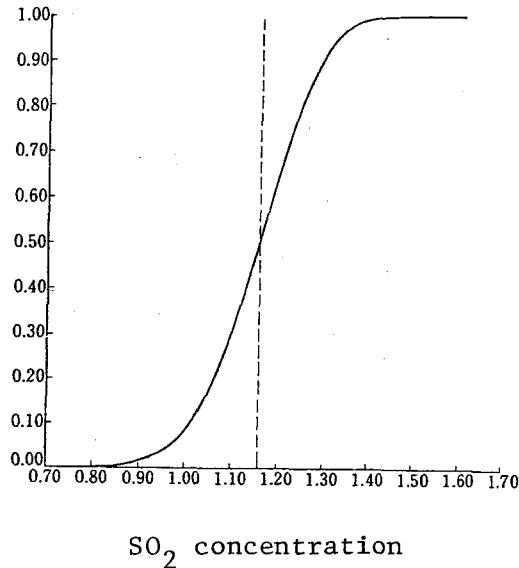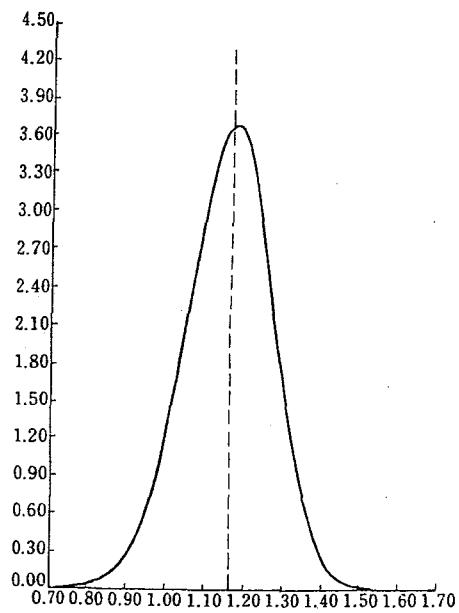
Fig. 2.3   Distribution function.

Fig. 2.4   Probability density function.

The distribution of $\hat{x}_o$ is nearly equal to a normal distribution and unimodal, but it is not symmetric and its tails are heavier than in a normal case.

The lower confidence limit of x_o, $\hat{x}_L$, can be obtained by the inverse estimation of a regression curve, for example Takeuchi [1], as

$$(2.11) \quad \hat{x}_L = \frac{(|\Sigma^{-1}|(\hat{\beta}_0 - \hat{\beta}_1)\hat{\beta}_2 - \frac{N_1 N_2}{N_1 + N_2} \bar{x}_2 \delta^2) - \sqrt{\delta^2 |\Sigma^{-1}|(\hat{\beta}'\Sigma^{-1}\hat{\beta} - \delta^2)}}{|\Sigma^{-1}|\hat{\beta}_2^2 - \frac{N_1 N_2}{N_1 + N_2} \delta^2},$$

where

$$\Sigma = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} (X_H{}'X_H)^{-1} \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}',$$

(2.12)

$$\delta^2 = F_{1,N_1+N_2-3;2\alpha} S_1/(N_1 + N_2 - 3),$$

$$S_1 = (y - X_H\hat{\beta})'(y - X_H\hat{\beta}),$$

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)'.$$

For various assurance levels, values of the lower confidence limits are given in Table 2.3.

Table 2.3 Lower confidence limits under the HS model with various kinds of assurance levels for data of Table 2.1.

Level α	Lower confidence limit
0.5	1.160
0.05	0.953
0.01	0.834
0.005	0.786

(2) General case

The data in Table 2.2 can not be separated into two groups. In this case, the estimator $\hat{x}_o$ is obtained by minimizing the residual sum of squares of the whole HS regression line.

Suppose that the pair data $(x_1, y_1), \ldots, (x_N, y_N)$ are arranged in the order of increasing magnitude in x. For each integer k ($k = 1, \ldots, N-1$), we give the HS regression model as follows.

(2.13)
$$y = x_H^k(\beta_0, \beta_1, \beta_2)' + e,$$

where
$$x_H^k = \begin{pmatrix} 1 & . & . & 1 & 0 & & . & . & 0 \\ 0 & . & . & 0 & 1 & & . & . & 1 \\ 0 & . & . & 0 & x_{k+1} & . & . & x_N \end{pmatrix}' \quad \text{and}$$

(2.14)
$$x_k \leqq \frac{\beta_0 - \beta_1}{\beta_2} < x_{k+1}.$$

We compute the least square error under the restriction (2.14) for each k, and seek the minimum value among them, $N-1$ in number. Let $\hat{k}$ and $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2)'$ be the estimators of k and $\beta = (\beta_0, \beta_1, \beta_2)'$, respectively. Then,

(2.15)
$$(y - x_H^{\hat{k}} \hat{\beta})'(y - x_H^{\hat{k}} \hat{\beta}) = \underset{1 \leqq k \leqq N-1}{Min} \quad \underset{x_k \leqq \frac{\beta_0-\beta_1}{\beta_2} < x_{k+1}}{Min} \quad (y - x_H^k \beta)'(y - x_H^k \beta).$$

Minimization can be solved by a fairly easy computation (Hudson [7], Hasselblad et al. [10]). The estimator of x_o is given by

(2.16)
$$\hat{x}_o = \frac{\hat{\beta}_0 - \hat{\beta}_1}{\hat{\beta}_2}.$$

The exact distribution of $\hat{x}_o$ has not been given because of the restriction (2.14). Therefore, the exact lower confidence limit is not given. Hasselblad et al. [10] gave an approximate value by the asymptotic normality of $\hat{x}_o$ (Hinkley [8]). For Table 2.2, the original authors [11] gave $\hat{x}_L = 0.133$

as the approximate 95 percent confidence lower limit. But, in a epidemiological or environmental study, asymptotic properties may not be reliable because the sample size is restricted to maintain the homogeneity in the data. For example, $N = 9$ in Table 2.2. More detailed research for the confidence interval is desirable.

2.1.3 Evaluation of the HS Model

The HS regression model is based on two assumptions. One is for the regression curve, and the other is for the error terms. We shall first consider the latter.

Though the HS model employs the normality for the error terms, their true distribution is the binomial distribution. Let n_i and y_i denote the sample size and the sample prevalence rate, respectively, at a dose level x_i ($i = 1, \ldots, N$). Then

$$(2.17) \qquad n_i y_i \quad \sim \quad Bi(n_i, p(x_i)).$$

If n_i is large, y_i has an asymptotically normal distribution, that is,

$$(2.18) \qquad y_i \quad \sim \quad N(p(x_i), \frac{p(x_i)(1 - p(x_i))}{n_i}).$$

Thus, the assumption of normality holds asymptotically but the assertion of the homogeneity of variances contradicts the expression $p(x_i)(1 - p(x_i))/n_i$ in (2.18).

Suppose that the population prevalence rate at each survey point is a random variable denoted by

(2.19) $$p(x, e) = p(x) + e,$$

where e represents the deviation from $p(x)$, such as area variation or survey error. In contrast with designed experiments, epidemiological surveys are not sufficient to control data, so that the population prevalence rate is not represented by x only. From (2.18) and (2.19), the variance of y_i is determined as follows.

(2.20) $$V(y_i) = p(x_i)(1 - p(x_i))/n_i + V(e)(1 - 1/n_i).$$

The first term is the sampling error and the second is the individual deviation. The assumption of (2.19) enlarges the variance of (2.18). If n_i and $p(x_i)$ are nearly homogeneous and the individual deviation exists, the normality with homogeneous variances holds approximately.

In Table 2.1, the mean square error is computed as 0.130×10^{-4}. On the other hand, if we put $p = 0.03$, $n = 2000$, we have $p(1 - p)/n = 0.145 \times 10^{-4}$. These two values are nearly equal. Thus the error of Table 2.1 data may be explained by the sampling error, that is, the binomial error. We may neglect the individual deviation. It shows the good quality of the data. We are able to make a similar discussion on Table 2.2.

Next, we shall evaluate the assumption of the regression curve. Though the HS regression line is simple, the assumption of the existence of the threshold value must be fulfilled. To avoid this difficulty, we may use

a smoothly increasing regression curve. One of most popular regression

curves to interpret a dose-response relationship is the probit curve

denoted by

(2.21) $p(x) = \beta_0 + (1 - \beta_0) \Phi(\beta_1 + \beta_2 \log(x))$,

where $\Phi(x)$ is the distribution function of the standard normal distribution

and β_0 means a spontaneous prevalence rate which is assumed to be positive.

Under the binomial model of (2.17), we shall compare the HS regression

line with the probit curve. The estimation method of the probit curve is

the well known probit analysis [9]. The results for Table 2.1 and Table

2.2 are described in Fig. 2.5 and Fig. 2.6, respectively. The fit of

the curve in each case is good.
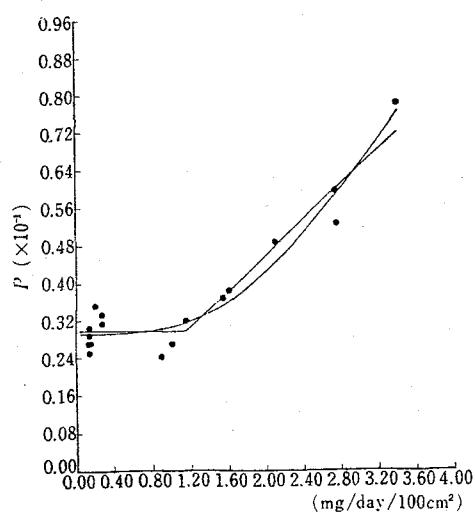

Fig. 2.5 Fitted regression
   lines of Table 2.1.
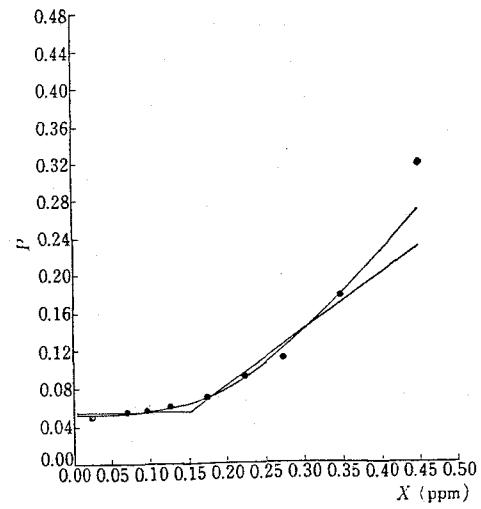
Fig. 2.6 Fitted regression
          lines of Table 2.2.

In Table 2.1, the chi-square value with 14 degrees of freedom for the test of goodness of fit is 9.138 under the HS model and 4.761 under the probit model. The fit is better for the probit curve than for the HS regression line.

In Table 2.2, the chi-square value with 6 degrees of freedom for the test of goodness of fit is 21.534 under the HS model and 6.436 under the probit model. The former is statistically significant with level 1 %. As the figure shows, the probit curve fits better.

This discussion shows that the probit model is suitable for Table 2.1 and Table 2.2. Since the probit model has no threshold value, we need a new definition of safety concentration instead of a threshold value. It is done by introducing a risk level. We discuss it in the next section.

2.2 PROBIT ANALYSIS AND A RISK LEVEL

In Section 2.1, we have a critical review of the HS regression method. The assumptions of the regression curve and the homogeneity of variances are not consistent. The fit of the probit curve is superior to that of the HS regression line.

Now we shall discuss the safe dose in the probit model. We introduce a risk level ρ which can be afforded from a social point of view. Then we define a safe dose x_o by an introduced risk level ρ in the probit curve of (2.21).

$$(2.22) \qquad \beta_0 + (1 - \beta_0)\Phi(\beta_1 + \beta_2 log(x_o)) = \beta_0 + (1 - \beta_0)\rho,$$

that is,

$$(2.23) \qquad log(x_o) = (\Phi^{-1}(\rho) - \beta_1)/\beta_2.$$

This definition is in line with Mantel and Bryan [5] and others, who presented methods to estimate the safe dose against carcinogenicity from experimental data.

An estimator $\hat{x}_o$ of the safe dose x_o is defined by a lower confidence limit of LD_ρ, the lethal dose at level ρ, with an assurance level $1 - \alpha$, which can be obtained by a well known technique in the probit analysis [9].

Here we choose five values, 0.01, 0.005, 0.001, 0.0005 and 0.0001, as risk levels. A very small value like 10^{-8} was adopted as a risk level by Mantel and Bryan [5] to estimate the safe dose against carcinogenicity, a fatal disease. But we do not choose such a small risk level, since chronic bronchitis and eye discomfort are not serious disease, but may be only a symptom.

For the data shown in Table 2.1, the fitted curve is given by

$$(2.24) \qquad 0.0289 + (1 - 0.0289)\Phi(-2.917 + 2.377log(x))$$

which is described in Fig. 2.5. The proposed value $\hat{x}_o$ for several risk levels and assurance levels are given in Table 2.4. Table 2.3 and Table 2.4 show that values of $\hat{x}_o$ obtained by the HS regression method are approximately equal to values of $\hat{x}_o$ obtained by the probit method with a risk level about 0.005 or less.

Table 2.4 Lower confidence limits under the probit model with various kinds of risk and assurance levels for the data of Table 2.1.

Assurance level	Risk level ρ				
α	0.01	0.005	0.001	0.0005	0.0001
0.5	1.780	1.398	0.849	0.700	0.462
0.05	1.461	1.055	0.537	0.412	0.234
0.01	1.286	0.877	0.396	0.290	0.149
0.005	1.213	0.806	0.344	0.247	0.121

In Table 2.2, the maximum likelihood estimates of $(\beta_0, \beta_1, \beta_2)$ are given by

$$(2.25) \qquad 0.0523 + (1 - 0.0523)\Phi(0.402 + 3.307log(x)).$$

It is described in Fig. 2.6. Lower confidence limits with several assurance levels and risk levels are given in Table 2.5. Lower confidence limits with an assurance level 0.5 are reduced to point estimates. The estimate $\hat{x}_o$ under the HS regression method is 0.145. The author does not calculate confidence intervals, since the test of goodness of fit of the HS regression line is statistically significant.

Table 2.5 Lower confidence limits under the probit model with various kinds of risk and assurance levels for the data of Table 2.2.

Assurance level	Risk level ρ				
α	0.01	0.005	0.001	0.0005	0.0001
0.5	0.150	0.126	0.088	0.076	0.057
0.05	0.134	0.110	0.073	0.062	0.044
0.01	0.127	0.103	0.066	0.055	0.038
0.005	0.124	0.100	0.063	0.053	0.036

2.3 DISCUSSION

Some conclusions and suggestions can be given through the above applications and other experiences.

The HS regression method is of omnibus use. In fact, the model is often well fitted rather than a simple linear regression model. But a defect is lack of scientific and medical interpretations of a safe dose x_o. It is hoped to be determined as the intersection of both lines. But for this purpose we need a certain physiological proof. That is, it is necessary to show the existence of a threshold value. Practically the model is often assumed only for convenience. A statistician is usually convinced that the dose-response curve is smoothly increasing, even when he uses the HS regression method.

A model with a smoothly increasing regression curve can delete this serious problem, but brings another difficulty. The regression curve does not present a point which suggests a safe dose directly. Thus a risk level is introduced to define a safe dose. This definition may be more natural than that by an intersection in the HS regression line.

Another trouble is how we choose a suitable family of regression curves. Fortunately, we have many conventional models of dose-response relationship, for example, the probit model, the logistic model, and so on. Our two examples show that the probit model is well fitted, even though the data are obtained not from designed experiments but from observational surveys.

The polynomial regression models are frequently used, when the linear regression model is not well fitted, but they are not applicable to

our problems. In fact, the regression model using the polynomial of order 3 is well fitted to both data, but the estimated regression curves are unacceptable because they are decreasing locally in the range of data.

REFERENCES

[1] 竹内啓（1966）。 数理統計学、東洋経済新報社。

[2] 高田目啓ら（1970）。 赤穂市における大気汚染と慢性気管支炎について、兵庫県公害研究所報告、第1号、25－34。

[3] 日本規格協会（1972）。 統計数値表。

[4] Fieller, E. C. (1932). The distribution of the index in a normal bivariate population, *Biometrika*, Vol. 24, pp. 428-441.

[5] Mantel, N. and Bryan, W. R. (1961). "Safety" testing of carcinogenic agents, *Jour. Nat. Cancer Inst.*, Vol. 27, pp. 455-470.

[6] British Medical Research Council (1965). Definition and classification of chronic bronchitis for clinical and epidemiological purposes, *Lancet*, Vol. 7389, pp. 755-779.

[7] Hudson, D. J. (1966). Fitting segmented curves whose join points have to be estimated, *J. Amer. Statist. Ass.*, Vol. 61, pp. 1097-1129.

[8] Hinkley, D. V. (1969). Inference about the intersection in two-phase regression, *Biometrika*, Vol. 56, pp. 495-504.

[9] Finney, D. J. (1971). *Probit analysis (third edition)*, Cambridge University Press.

[10] Hasselblad, V., et al. (1973).   Regression using "hockey stick"

functions, Paper presented at the Statistical Section, 101st

Annual Meeting, Am. Public Health Ass., San Francisco, Nov. 8, 1973.

[11] Hummer, D. I., et al. (1974).   Los Angeles student nurse study

"Daily symptom reporting and photochemical oxidants", *Arch. Environ.*

Health, Vol. 28, pp. 255-260.

3 Conversion of a Safe Dose: Critical Review of

the Environmental Standard of Nitrogen Dioxide

The environmental standard of NO_2 is based on the criterion which is
obtained from the results of biological experiments for acute and chronic
toxicity and epidemiological surveys for chronic bronchitis etc. There
are two standards of NO_2 in Japan. One is the short-range standard which
gives the upper limit of an hour mean of NO_2 concentrations for the
protection against acute toxicity. The other is the long-range standard
which gives the upper limit of a day mean of NO_2 concentrations for
the protection against chronic toxicity.

The environmental standards are used by national or local administrations
for daily monitoring and controlling air pollutions. In air polluted
areas, plans of controlling the pollutant sources to satisfy standards
are drawn and performed. When a new air pollutant discharge source is
built, the assessment for obeying the standards must be done.

In 1978, the Environmental Agency of Japan [7] changed the long-range
environmental standard of NO_2 from "the rate of days which satisfy
the standard of 0.02 ppm/day to all days in a year must be more
than 98 percent" into "the rate of days which satisfy the standard of

0.04 ∿ 0.06 ppm/day to all days in a year must be more than 98 percent".

This change raised many problems and was criticized by many scientists:

[8] among others. For example, the safe coefficient of 1/2 which had

been adopted in the previous standard [1], [2] was discarded in the new

one.

Apart from these criticism, the author questioned whether the conversion

method which was used to obtain the long-range standard is appropriate

or not. In the announcement [7] by the Environmental Agency, the part

related to the conversion is read as follows. "Though the report [6]

gave the criteria, that is, an hour mean 0.1 ∿ 0.2 ppm against short-

range exposure and a year mean 0.02 ∿ 0.03 ppm against long-range exposure,

the new long-range environmental standard would be given by a day mean

as well as the previous standard. Since there exists a close relationship

between day means and year means, the concentration 0.04 ∿0.06 ppm day

mean is nearly equivalent to the concentration 0.02 ∿ 0.03 ppm year mean,"

where the day mean actually stands for the 98 percent value of day means

in a year, which will be denoted by the 98 % value. The conversion from

a year mean to a 98 % value was based on the linear regression function.

The linear regression function used by the Environmental Agency is the

one which was computed from NO_2 data observed at the stations, 1114 in

number, of measuring air pollutions throughout Japan during three years,

1973 ∿ 1975. In the Report [5], the linear regression function of 98 %

values $x(0.98)$ on year means m is reported to be

$$x(0.98) = 1.82m + 0.0045 \quad (ppm),$$

correlation coefficient = 0.919.

With this relation, a year mean 0.02 ppm is converted to a 98 % value

0.04 ppm.

But, it is dangerous that a linear regression function is used for

the conversion solely on the ground of high correlation. Evaluation for

the error terms in the regression model should be made. Since the

environmental standard plays an important role in society, a statistical

analysis to the above conversion should be performed.

In the following discussion, we put the assumption of the lognormal

model for NO_2 concentrations. Under the assumption, we may introduce

the conversion method which uses the linear regression function with

the error term whose standard deviation is proportional to the independent

variable and which adopts the tolerance limits. Since the lognormal

model for NO_2 concentrations is not well fitted in some areas, robustness

of the conversion method is discussed. NO_2 data in 1977 at Okayama

Prefecture are used for illustration throughout the following statistical

analysis.

3.1 THE CURRENT CONVERSION METHOD

3.1.1 Review of the Current Conversion Method

The current conversion method uses the linear regression function of

98 % values on year means. We shall apply it to NO_2 data in Okayama 1977.

Let N be the number of stations of measuring NO_2. Let x_{ij} $(j = 1, \ldots, n_i)$

be available measurments of day means at the ith station in a year.

The year mean m_i and the 98 % value $x_i(0.98)$ of the ith station are estimated respectively by

(3.1)
$$\hat{m}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij},$$

(3.2)
$$\hat{x}_i(0.98) = x_{i([0.98n_i + 1])},$$

where [ ] is the Gauss symbol and $x_{i(j)}$ is the jth order statistic in the ith sample. From the estimated year means and 98 % values of 46 stations in Okayama 1977, we get the following estimated linear regression function, which is described in Fig. 3.1.

(3.3)
$$x(0.98) = 1.623\,m + 0.006,$$

$$\sqrt{\text{mean square error}} = 0.0093,$$

$$\text{correlation coefficient} = 0.926.$$

Using the equation (3.3), a year mean 0.02 is converted to a 98 % value 0.039. We shall discuss the conversion method in the following section.

x(0.98)
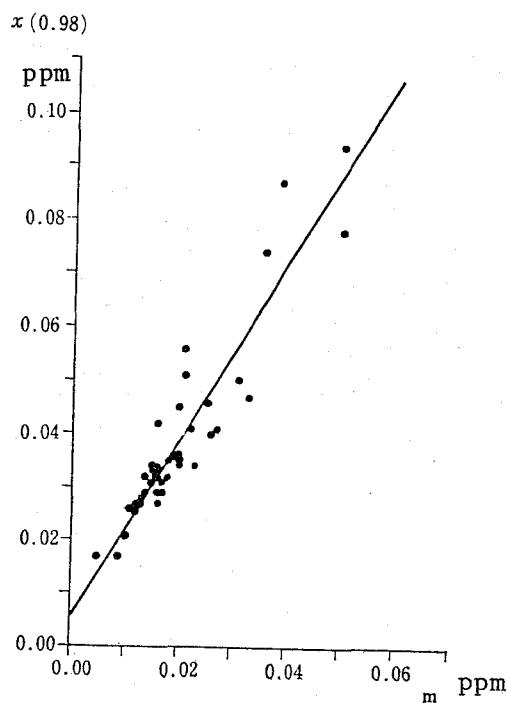
Fig. 3.1 Year means and 98 % values of NO_2 concentrations
at 46 stations in Okayama 1977.

3.1.2 Distribution of NO_2 Concentrations

As usual, the distribution of concentrations of an air pollutant
subject such as SO_2 or NO_2 is considered to fit the lognormal distribution
which has the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} exp[- \frac{1}{2\sigma^2} (log(x) - \mu)^2].$$

The parameter μ is thought to represent the degree of concentrations
and the parameter σ represent the area property such as meteorological
and geographical features and air pollutant subject discharge forms.

We examined the goodness of fit of the lognormal model to NO_2 data

in Okayama 1977. Two parameters (μ_i, σ_i^2) at the ith station ($i = 1,\ldots, N$)

were estimated by the following two estimators, respectively.

(3.4)
$$\hat{\mu}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} log(x_{ij}),$$

(3.5)
$$\hat{\sigma}_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (log(x_{ij}) - \hat{\mu}_i)^2.$$

For each station, a X^2 test of the goodness of fit was performed by

deviding concentration range into 10 intervals with equal probabilities.

The results showed that 20 stations among 46 were statistically significant

at 5 % level. NO_2 data of these stations plotted in the lognormal

probability papers showed the upward deviation from the straight line

in the lower concentrations. For example, lognormal plot of NO_2 data

of the Siose station is shown in Fig. 3.2. The value of X^2 statistic

of the Siose station was 47.7 with 7 degrees of freedom and statistically

significant at 1 % level. We remark that the accuracy of measurement

in the low level is worse and the measurments under a certain value

are recorded as zero.

The caracteristic values for the conversion such as year means and

98 % values are middle and high level values. Therefore, if we can

devise estimators without using low level data, we may escape from the

deviation of the lognormal model in the low level. Using the data
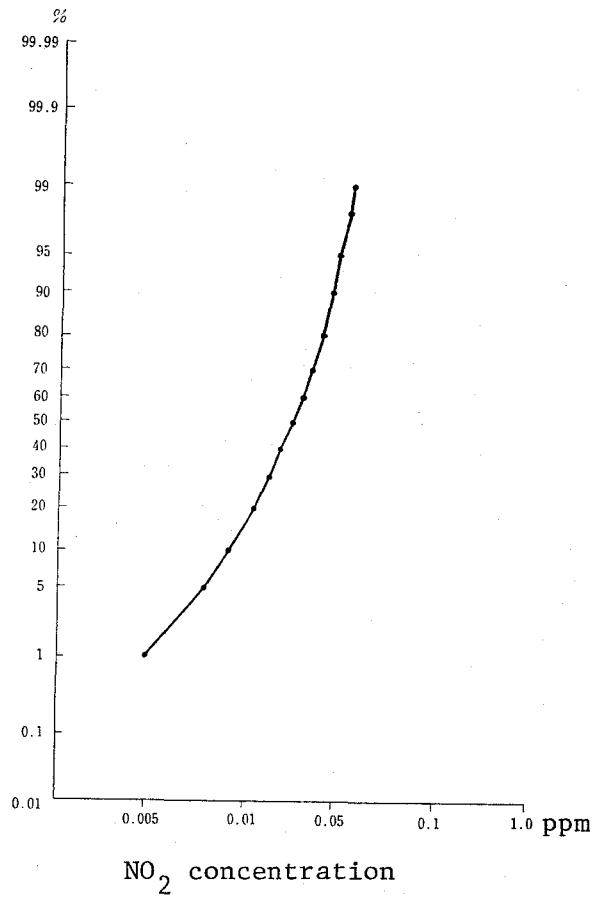
Fig. 3.2 NO_2 data of the Siose station plotted in the lognormal probability paper.

above the median, we estimated (μ_i, σ_i^2) $(i = 1,\ldots, N)$ by the following estimators.

(3.6)
$$\tilde{\mu}_i = log(x_{i([0.5 n_i + 1])}),$$

(3.7)
$$\tilde{\sigma}_i^2 = \sum_{log(x_{ij}) \geq \tilde{\mu}_1} (log(x_{ij}) - \tilde{\mu}_i)^2 / [(n_i + 1)/2].$$

The diagram of $(\tilde{\mu}_i, \tilde{\sigma}_i)$ ($i = 1, \ldots, 46$) computed from data in Okayama 1977

is shown in Fig. 3.3.
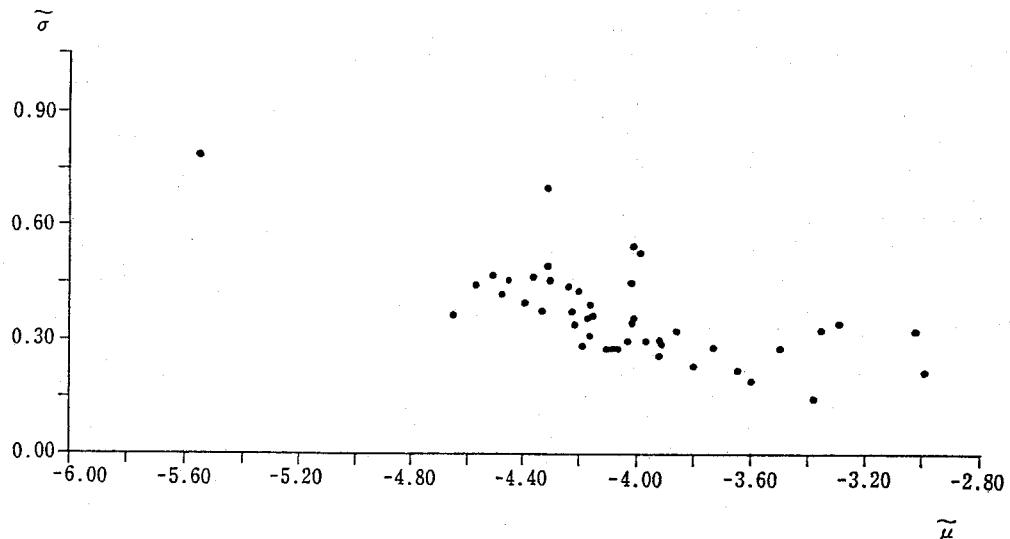
Fig. 3.3 Diagram of (μ, σ) in Okayama 1977.

The X^2 tests of the goodness of fit which pool the intervals below

the medians reduced the number of significant stations at 5 % level to

10. Here the pattern of NO_2 concentrations of significant stations

showed that the deviation from the straight line in the lognormal

probability paper tended below. One explanation will be that when

high concentrations appeared, the NO_2 discharge sources were controlled

so that high concentrations were restrained.

In the following discussion, we first assume the lognormal model

for NO_2 concentrations and use the estimators of (3.2), (3.6) and (3.7).

Under the model, we criticize the current conversion method and introduce

a new one. Then, the robustness of the new method is discussed.

3.1.3 Critical Discussion for the Current Conversion Method

Assume the lognormal model, that is, NO_2 concentrations (day means) x_{ij} ($j = 1,\ldots, n_i$) at the ith station are distributed lognormally with the parameters (μ_i, σ_i). Then, the year mean m_i and the 98 % value $x_i(0.98)$ are represented in terms of μ_i and σ_i as follows.

$$(3.8) \qquad m_i = exp(\mu_i + \frac{\sigma_i^2}{2}),$$

$$(3.9) \qquad x_i(0.98) = exp(\mu_i + \Phi^{-1}(0.98)\, \sigma_i),$$

where $\Phi^{-1}(\ )$ is the inverse function of the standard normal distribution function $\Phi(\ )$. Therefore, the following relation between m_i and $x_i(0.98)$ holds.

$$(3.10) \qquad x_i(0.98) = exp(\Phi^{-1}(0.98)\, \sigma_i - \frac{\sigma_i^2}{2})\, m_i.$$

It shows that if all σ_i's for N stations are equal, $x(0.98)$ and m have the proportional relationship.

From the above discussion, the current conversion method of using the linear regression function may be considered to have postulated two assumptions. The first is the homogeneity in σ_i's of the error terms in the linear regression model and the second is that the error terms can be explained as sampling errors only. Under the assumption of sampling errors, $N = 1114$ is so large that confidence intervals are meaningless, that is, they become to be equal to the point estimation.

But two assumptions do not hold in practice. As easily seen in Fig. 3.3, the homogeneity of σ_i's does not hold in Okayama 1977. Moreover, the deviations from the linear regression line are larger than sampling errors. It will be seen in the following argument.

From the theory of order statistics, $\hat{x}_i(0.98)$ is asymptotically distributed in the normal distribution

(3.11)
$$N\left(exp(\Phi^{-1}(0.98)\sigma_i - \frac{\sigma_i^2}{2})\,m_i,\ \frac{0.04\pi\sigma_i^2}{n_i}\ exp(\{\Phi^{-1}(0.98)\}^2 + 2\Phi^{-1}(0.98)\sigma_i - \sigma_i^2)m_i^2\right).$$

We shall examine whether the asymptotic normal distribution (3.11) can be used as an approximate distribution of $\hat{x}_i(0.98)$. Average numbers of days with available measurments of 46 stations in Okayama 1977 is 345. Since $[0.98 \times 345 + 1] = 339$, $\hat{x}(0.98)$ is the 7th largest order statistic. The average values of 46 stations are $n = 345$, $\mu = -4.10$ and $\sigma = 0.4206$. We compare the exact distribution of $\hat{x}(0.98)$ with the asymptotic distribution of (3.11) by using the parameter values computed above. Two probability density functions are described in Fig. 3.4. The exact distribution is not symmetric and skewed positively. It has the mean 0.0397, the standard deviation 0.00267 and the skewness 0.36. The asymptotic distribution (3.11) has the mean 0.0393 and the standard deviation 0.00257. The latter is 1 % smaller in the mean and 4 % smaller in the standard deviation in comparison with the former.

From the results, we may say that in the size of n_i now discussed, $\hat{x}(0.98)$ is distributed approximately with the mean and variance of the
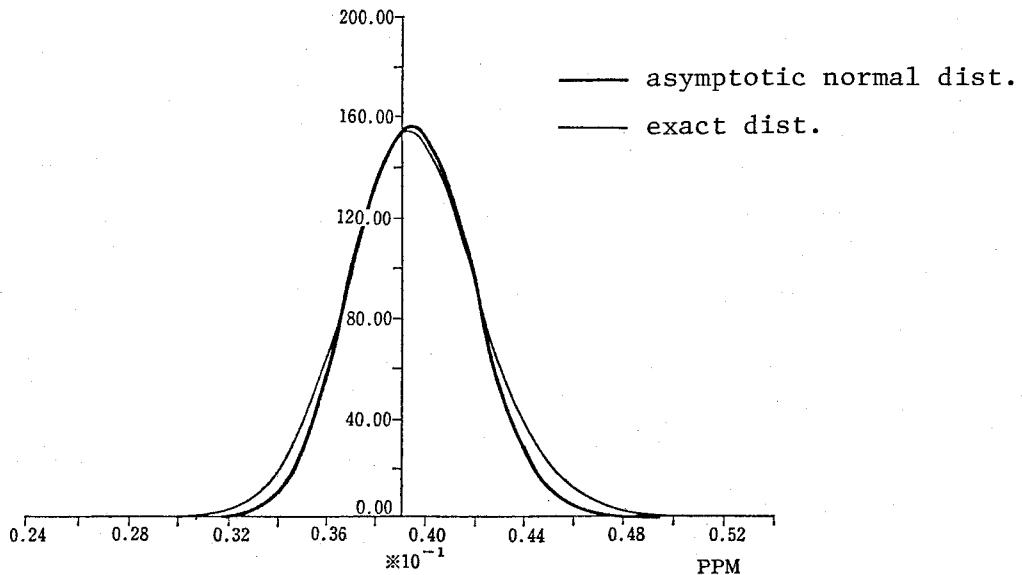
Fig. 3.4 Exact and asymptotic distributions of $\hat{x}(0.98)$.

normal distribution (3.11) and skewed positively. Therefore, the regression

line of $x_i(0.98)$ on m_i is approximately the regression line of (3.11)

with the standard deviation of sampling error proportional to m_i:

The standard deviation of sampling error will be computed. At the

typical station of Kencenter, it is 0.0032. It is smaller than the

square root of the mean square error of the regression line, 0.0093.

This results shows that the error term in the regression model is not

explained by the sampling error only. The σ_i varies among stations,

so that regression coefficient varies too. Let (k_1, k_2) be the range

of the regression coefficients for all stations, then the linear regression

line has the width of $(k_2 - k_1)m$. From the above discussion we may

conclude that the linear regression line must be considered to represent

the model of the average or typical station and to have the error term

whose standard deviation is proportional to m, that is, proportional to

$x(0.98)$, and which consists of sampling error and individual deviation

yielded from the variation of σ_i, that is, area variation.

3.2 THE AUTHOR'S CONVERSION METHOD

In this section, Using the linear regression model with the error

term whose standard deviation is proportional to the independent variable,

we shall introduce a conversion method of a year mean to a 98 % value.

Since our purpose is to control m by the converted value $x(0.98)$, the

linear regression of m on $x(0.98)$ will be reasonable. From the discussion

in the previous section, the following empirical model is assumed.

(3.12) $$M = \alpha\, x(0.98) + \beta + e,$$

$$e \sim N(0,\ \sigma_o^2\, x^2(0.98)),$$

where M is a random variable which represents a year mean and the error

term e consists of sampling error and individual deviation. Since we

can not separate the two, the conversion should be done by the tolerance

limit theory which can evaluate the individual deviation of area properties.

Introducing three levels; c: threshold level, ρ: risk level and γ: assurance

level, we define the converted value $\hat{x}_o$ which satisfies the following

relation.

(3.13) $$Pr\{Pr\{M \geq c \mid x(0.98) \leq \hat{x}_o\} \leq \rho\} \geq 1 - \gamma.$$

The definition means that it holds with probability at least $1 - \gamma$ that when $x(0.98)$ is under $\hat{x}_o$, the probability that the year mean is larger than the threshold level c is less than the risk level ρ.

Transforming M and $x(0.98)$ into

$$(3.14) \qquad z = \frac{M}{x(0.98)} \quad \text{and} \quad y = \frac{1}{x(0.98)},$$

we obtain the well known model

$$(3.15) \qquad z = \alpha + \beta y + w,$$

$$w \sim N(0, \sigma_o^2),$$

in place of (3.12). Using N transformed pair data, (y_i, z_i) $(i = 1, \ldots, N)$, we can estimate $(\alpha, \beta, \sigma_o)$ with a usual method. From Takeuchi [3], $\hat{x}_o$ can be computed with the upper tolerance limit $\hat{g}(y)$ of z in the following manner.

$$(3.16) \qquad Pr\{Pr\{Z \geq \hat{g}(y)\} \leq \rho\} \geq 1 - \gamma,$$

$$(3.17) \qquad \hat{x}_o \hat{g}\left(\frac{1}{\hat{x}_o}\right) = c,$$

where

$$(3.18) \quad \begin{cases} \hat{g}(y) = \hat{\alpha} + \hat{\beta}y + k(y)\hat{\sigma}_o, \\[2mm] k(y) = h(y)t_\gamma(n - 2, \lambda(y)), \\[2mm] h(y) = (\dfrac{1}{n} + \dfrac{(y - \bar{y})^2}{\displaystyle\sum_{i=1}^{N} (y_i - \bar{y})^2})^{\frac{1}{2}}, \\[4mm] \lambda(y) = U_\rho/h(y), \\[2mm] t_\gamma(n, \xi): \text{ upper } 100\gamma \ \% \text{ point of the non-central} \\ \qquad\quad \text{t distribution with } n \text{ degrees of freedom} \\ \qquad\quad \text{and non-centrality parameter } \xi, \\[2mm] U_\rho: \text{ upper } 100\rho \ \% \text{ point of the standard normal distribution.} \end{cases}$$

For the upper 100γ % point of the non-central t distribution, we use the following approximation:

$$(3.19) \qquad t_\gamma(n, \xi) = \frac{U_\gamma + \xi}{\sqrt{1 - U_\gamma^2/2n}},$$

then (3.17) can be rewritten as

$$(3.20)$$
$$\hat{\beta} + (\hat{\alpha} + \frac{U_\rho\hat{\sigma}_o}{\sqrt{1 - U_\gamma^2/2(n - 2)}})\hat{x}_o + \frac{U_\rho\hat{\sigma}_o}{\sqrt{1 - U_\gamma^2/2(n - 2)}}\sqrt{\frac{\hat{x}_o^2}{n} + \frac{(1 - \bar{y}\hat{x}_o)^2}{\displaystyle\sum_{i=1}^{N} (y_i - \bar{y})^2}} = c.$$

The value $\hat{x}_o$ can be obtained as the larger solution of the following quadratic equation,

$$(3.21) \qquad a_2 \hat{x}_o^2 + 2a_1 \hat{x}_o + a_0 = 0,$$

where

$$(3.22)$$

$$
\begin{cases}
a_0 = (c - \hat{\beta})^2 - \dfrac{U_\gamma^2 \hat{\sigma}_o^2}{(1 - U_\gamma^2/2(n - 2))(\sum\limits_{i}(y_i - \bar{y})^2)} , \\[4mm]
a_1 = \dfrac{U_\gamma^2 \hat{\sigma}_o^2 \bar{y}}{(1 - U_\gamma^2/2(n - 2))(\sum\limits_{i}(y_i - \bar{y})^2)} - (c - \hat{\beta})(\hat{\alpha} + \dfrac{U_\gamma \hat{\sigma}_o}{\sqrt{1 - U_\gamma^2/2(n - 2)}}), \\[4mm]
a_2 = (\hat{\alpha} + \dfrac{U_\rho \hat{\sigma}_o}{\sqrt{1 - U_\gamma^2/2(n - 2)}})^2 - \dfrac{U_\gamma^2 \hat{\sigma}_o^2}{1 - U_\gamma^2/2(n - 2)} (\dfrac{1}{n} + \dfrac{\bar{y}^2}{\sum\limits_{i}(y_i - \bar{y})^2}).
\end{cases}
$$

Then we have

$$(3.23) \qquad \hat{x}_o = (-a_1 + \sqrt{a_1^2 - a_2 a_0})/a_2.$$

The conversion with the data in Okayama 1977 is shown in Table 3.1.

Table 3.1   Conversion of a year mean 0.02 ppm to 98 % values of day means under various risk levels and assurance levels.

Assurance level	Risk level ρ		
α	0.01	0.005	0.001
0.01	0.0273	0.0266	0.0251
0.05	0.0270	0.0263	0.0249
0.01	0.0265	0.0258	0.0244

(Threshold level $c = 0.02$)

3.3 DISCUSSION

The proposed conversion method is based on the assumption that the
data are random samples and the distribution of NO_2 concentrations is
lognormal. The assumption of randomness holds approximately, because
day means are grasped in a packege of one year so that time dependency
as for hour measurments is inconceivable.

As for the assumption of lognormal model, the badness of fit in low
level data can be delated by using the estimators (3.6) and (3.7).
The deviation of the lognormal model in high level data will be evaluated
later and we shall see the robustness of the new conversion method.

We evaluate the effect of the deviation against the proportional
relationship between year means and 98 % values in (3.10). Let $f(x)$
denote the real probability density function of NO_2 concentrations,
and let μ and σ denote the location and the scale parameters, respectively,
of log-transformed NO_2 concentrations, $log(x)$. Let $g(y)$ denote the real
probability density function of $y = (log(x) - \mu)/\sigma$. Since $f(x)$ may
be near by the lognormal distribution, $g(y)$ may be near by the standard
normal distribution.

Now let us expand $g(y)$ with the Gram-Charlier-Edgeworth expansion [4],
that is,

$$(3.24) \qquad g(y) = \phi(y)\{a_0 + a_1 h_1(y) + a_2 h_2(y) + \cdots\},$$

where $a_j = E_g[h_j(Y)/j!]$, $h_j(y)$ denotes the jth Hermite polynomial and
$\phi(\ )$ denotes the probability density function of the standard normal

distribution. Since we estimate (μ, σ) with the upper half data, $g(y)$ can be taken to be symmetric. Then the third moment, skewness, is zero.

We evaluate the deviation from the lognormal model by using the fourth moment μ_4, kurtosis. Thus,

(3.25) $a_0 = 1, \ a_1 = a_2 = a_3 = 0, \ a_4 = (\mu_4/\sigma^4 - 3)/24,$

(3.26) $g(y) = \phi(y) \ (1 + a_4 h_4(y) \ ).$

From (3.26), we evaluate the effect of kurtosis to the year mean m.

(3.27) $m = \displaystyle\int_0^\infty x f(x) \, dx$

$= \displaystyle\int_{-\infty}^\infty exp(\mu + \sigma y) g(y) \, dy$

$= \displaystyle\int_{-\infty}^\infty exp(\mu + \sigma y) \phi(y)(1 + a_4 h_4(y)) \, dy \qquad (y = z + \sigma)$

$= exp(\mu + \dfrac{\sigma^2}{2}) \displaystyle\int_{-\infty}^\infty \phi(z)[1 + a_4(3 - 6\sigma^2 + \sigma^4) + a_4(4\sigma^3 - 12\sigma)z$

$\qquad\qquad + a_4(6\sigma^2 - 6)z^2 + a_4 4\sigma z^3 + a_4 z^4] \, dz$

$= exp(\mu + \dfrac{\sigma^2}{2})(1 + a_4 \sigma^4).$

Since the distribution function $G(y)$ of Y is approximately equal to

(3.28) $G(y) = \Phi(y) - \phi(y) a_4 h_3(y),$

the 98 % point of Y denoted by y' can be represented by the sum of the 98 % point of the standard normal distribution denoted by y and a modification term $B(y)$ such that

$$(3.29) \qquad G(y') = \Phi(y) = 0.98,$$

$$(3.30) \qquad y' = y + B(y).$$

Because $B(y)$ can be considered to be small, we neglect the second order and have the following equation.

$$(3.31) \quad \Phi(y) = G(y') = G(y + B(y))$$

$$= \Phi(y + B(y)) - \phi(y + B(y))a_4h_3(y + B(y))$$

$$= \Phi(y) + \phi(y)[B(y) + a_4(y^4 - 6y^2 + 3)B(y) - a_4(y^3 - 3y)].$$

From (3.31), we have

$$(3.32) \qquad B(y) = \frac{a_4(y^3 - 3y)}{1 + a_4(y^4 - 6y^2 + 3)}.$$

Therefore, the effect of kurtosis to $x(0.98)$ is as follows.

$$(3.33) \qquad x(0.98) = exp(\mu + \sigma[\Phi^{-1}(0.98) + B(\Phi^{-1}(0.98))]).$$

From (3.27) and (3.33), we have

$$(3.34) \quad \frac{x(0.98)}{m} = exp(\sigma\Phi^{-1}(0.98) - \frac{\sigma^2}{2}) \frac{exp[\sigma B(\Phi^{-1}(0.98))]}{1 + a_4\sigma^4} .$$

The effect of kurtosis to the linear relationship between $x(0.98)$ and

m under fixed σ is measured by

$$(3.35) \quad 1 - \frac{exp[\sigma B(\Phi^{-1}(0.98))]}{1 + a_4\sigma^4} .$$

If (3.35) is neary zero, we can neglect the effect of kurtosis. We shall

check it for the NO_2 data in Okayama 1977. The mean and the standard

deviation of (3.35) for 46 stations are 0.009 and 0.024, respectively.

The maximum value and the mimimum value are 0.060 and -0.093, respectively.

Thus we can say that the effect of kurtosis is negligible so that the

new linear regression model is robust.

The above discussion shows that the current conversion method is not

appropriate because it provides no evaluation of the variability of the

area properties. Especially, the lack of the principle of safety coefficient

which may cover the above difficulty is serious.

Moreover, converted values in Table 3.1 should be considered as a

criterion because the year mean previously given is a criterion. Since

a criterion is given from a medical standpoint while a standard is given

from a administrative standpoint, a converted value can not be used directly

as a standard.

REFERENCES

[1]   中央公害対策審議会大気部会（１９７２）。　窒素酸化物等に係る環境基準についての
      専門委員会報告。

[2]   環境庁（１９７２）。　NO_2の環境基準と規制について。

[3]   竹内啓（１９７３）。　許容規準の定め方、応用統計学、第３巻、１－１３。

[4]   竹内啓（１９７５）。　確率分布の近似、教育出版。

[5]   環境庁大気保全局企画課（１９７７）。　二酸化窒素による我国の大気汚染の
      現況についての考察、公害と対策、第１３巻、７１１－７１５。

[6]   中央公害対策審議会（１９７８）。　二酸化窒素の人の健康影響に係る判定条件等
      について（答申）。

[7]   環境庁（１９７８）。　二酸化窒素に係る環境基準の改定について。

[8]   塚谷恒雄（１９７９）。　NO_2の有害性、公害研究、第９巻、２６－３０。

4 Dose-Response Relationship in Stratified Populations

Suppose that the prevalence rate of a disease and the intensity of a study factor that may be considered as the cause of the disease are observed in several districts, N in number. Let p_j be the prevalence rate and x_j the study factor intensity in the jth district. A X^2 test procedure for a linear regression of p_j on x_j, $p_j = \alpha + \beta x_j$ ($j = 1, \ldots, N$), was proposed by Cochran [3] and Armitage [4]. Their set of three X^2 tests for total differences, linear trend and slope in the prevalence rates has been frequently used in epidemiological survey works.

When some confounding factors, not the study factor, that may affect the prevalence rates of the disease exist, a X^2 test for the effect of the study factor on the prevalence rates has been done in the model of multiple $2 \times N$ contingency tables (c.f. Mantel[5], Birch [6], [7], Armitage [8]). Denote it as X^2_M-test. Recently, Wood [11] proposed other X^2 tests that are constructed by summing up X^2 statistics for the measure of fitness of linear models in each stratum. Neither of these procedures provides a set of three X^2 tests for total differences, linear trend, and slope in the prevalence rates, nor do they provide the estimators of the adjusted prevalence rates that can indicate the dose-response

relationship.

In Section 4.1, we propose a set of three X^2 tests for total differences, linear trend, and slope in the adjusted prevalence rates and then give their approximate powers. The X^2 test procedure involves controlling the confouding factors by stratification and adjusting the prevalence rates for the stratification, and is similar to the Cochran and Armitage procedure. Our tests are compared with the X_M^2-test and the X^2 tests performed by summing up X^2 statistics in each stratum in the style of Wood in Section 4.2. Their application to the data by Tsubota [1], [2] involving cough and phlegm is given in Section 4.3. Discussion is given in Section 4.4.

4.1 AN ESTIMATION AND TEST PROCEDURE FOR LINEAR TREND IN ADJUSTED PREVALENCE RATES

4.1.1 A X^2 Test Procedure

Let g be the number of the confounding factors and h_s be the number of levels of the sth confounding factor. Put $k = \prod\limits_{s=1}^{g} h_s$. To control the confounding factors, we subdivide each district into k strata. Suppose that the ith stratum of the jth district has w_{ij} population rate, $\sum\limits_{i} w_{ij} = 1$, and p_{ij} prevalence rate ($i = 1,\ldots, k$; $j = 1,\ldots, N$). Let n_{ij} be the number of observations and m_{ij} the number of cases with positive symptoms in the (i, j) cell. Let x_j ($j = 1,\ldots, N$) be the intensity of the study factor in the jth district.

Generally, the crude prevalence rate of each district, $p_j^* = \sum_{i=1}^{k} w_{ij} p_{ij}$
($j = 1, \ldots, N$), is not controlled with respect to the effects of the

confounding factors because the population composition varies among districts.

Thus we consider the adjusted prevalence rates defined by $p_j = \sum_{i=1}^{k} w_i p_{ij}$,

where $\{w_i, i = 1, \ldots, k; \sum_i w_i = 1, w_i > 0\}$ are the standard population

rates of the standard district.

Let us consider the hypotheses similar to those of Cochran and Armitage.

We denote by P_1 the set of the N-dimensional vectors $p = (p_1, \ldots, p_N)'$,

where $0 < p_j < 1$ ($j = 1, \ldots, N$) and then define two subsets of P_1,

$P_2 \doteq \{p; p = \alpha 1 + \beta x \in P_1, \alpha, \beta \in R^1\}$ and $P_3 \doteq \{p; p = \alpha 1 \in P_1, \alpha \in R^1\}$,

where $1 = (1, \ldots, 1)'$ and R^1 is the one dimensional Euclidian space

and $x = (x_1, \ldots, x_N)'$. We denote in general by $H(P_i; P_j)$ the hypothesis

H that $p \in P_i$ against the alternative $p \in P_j - P_i$. We say that H is

true if $p \in P_i$ and false if $p \in P_j - P_i$. The assertion that $p \in P_i$ is

referred to as the null hypothesis of H. The 1st hypothesis $H_1(P_3; P_1)$

is for total differences in the adjusted prevalence rates. The 2nd

hypothesis $H_2(P_2; P_1)$ is for linear trend. The 3rd hypothesis $H_3(P_3; P_2)$

is for the slope of the linear trend.

Now, we shall give a X^2 test procedure for $H_1(P_3; P_1)$, $H_2(P_2; P_1)$

and $H_3(P_3; P_2)$. If we have exact information of the confounding factors

in each district before sampling, we can stratify each district into

k strata by using this information and select a sample of appropriate

size from each stratum. But, in many epidemiological survey, we have

no such information before sampling. In these cases, we select a sample

of size $n._j$ from the jth district, and stratify it into k strata by

using the subdata of confounding factors of observations. Since w_{ij} is a nuisance parameter and its sufficient statistic is $n_{ij}/n_{.j}$, we shall consider the conditional tests given $\{n_{ij}\}$.

We estimate p_{ij} by $\hat{p}_{ij} = m_{ij}/n_{ij}$, where m_{ij} has the binomial distribution $Bi(n_{ij}, p_{ij})$. Using the prior standard population rates, we estimate p_j by $\hat{p}_j = \sum_i w_i p_{ij}$.

If a sequence of random variables $\{X_n\}$ converges in probability to a constant c, we denote it by $X_n \xrightarrow{P} c$. If $\{X_n\}$ converges in law to the distribution D, we denote it by $X_n \xrightarrow{\ell} D$.

When each n_{ij} tends to infinity subject to $n_{ij}/n_{..}$ remaining fixed, it holds that

(4.1)
$$\frac{\sqrt{n_{ij}}(\hat{p}_{ij} - p_{ij})}{\sqrt{p_{ij}(1 - p_{ij})}} \xrightarrow{\ell} N(0, 1) \qquad \text{and}$$

(4.2)
$$\frac{(\hat{p}_j - \sum_i w_i p_{ij})}{\sqrt{\sum_i w_i^2 \dfrac{p_{ij}(1 - p_{ij})}{n_{ij}}}} \xrightarrow{\ell} N(0, 1).$$

We define the jth adjusted sample size n_{oj} by

$$n_{oj} = \frac{(\sum_i w_i p_{ij})(1 - \sum_i w_i p_{ij})}{\sum_i w_i^2 \dfrac{p_{ij}(1 - p_{ij})}{n_{ij}}},$$

which is an increasing function of $\{n_{ij}, \ i = 1, \ldots, k\}$. Using n_{oj}, we can simplify (4.2) to a well known expression:

$$(4.3) \qquad \frac{\sqrt{n_{oj}}(\hat{p}_j - p_j)}{\sqrt{p_j(1 - p_j)}} \xrightarrow{\ell} N(0, 1).$$

By using the weights, $\omega_j = n_{oj} / p_j(1 - p_j)$ $(j = 1, \ldots, N)$, we can give a set of X^2 test statistics for $H_1(P_3; \ P_1)$, $H_2(P_2; \ P_1)$ and $H_3(P_3; \ P_2)$ which satisfy the following identity:

$$(4.4) \qquad \sum_j \omega_j(\hat{p}_j - \hat{p})^2 = \sum_j \omega_j(\hat{p}_j - \tilde{p}_j)^2 + \sum_j \omega_j(\tilde{p}_j - \hat{p})^2,$$

where $\hat{p} = \sum_j \omega_j \hat{p}_j / \sum_j \omega_j$, $\tilde{p}_j = a + bx_j$, $a = \hat{p} - b\bar{x}$,

$$b = \sum_j \omega_j \hat{p}_j(x_j - \bar{x}) / \sum_j \omega_j(x_j - \bar{x})^2 \text{ and } \bar{x} = \sum_j \omega_j x_j / \sum_j \omega_j.$$

From (4.3) and normal regression theory, it holds that the 1st X^2 test statistic is asymptotically distributed according to the X^2 distribution with $N - 1$ degrees of freedom, which will be denoted by X^2_{N-1} in the sequel, if H_1 is true. The 2nd X^2 test statistic is similarly asymptotically distributed according to X^2_{N-2} if H_2 is true, and the 3rd X^2 test statistic is asymptotically distributed according to X^2_1 if H_3 is true.

The parameters ω_j are unknown, so that we estimate them by

$$\hat{\omega}_j = \frac{\hat{n}_{oj}}{\hat{p}_j(1 - \hat{p}_j)} \qquad (j = 1, \ldots, N),$$

where

$$\hat{n}_{oj} = \frac{(\sum_i w_i \hat{p}_{ij})(1 - \sum_i w_i \hat{p}_{ij})}{\sum_i w_i^2 \dfrac{\hat{p}_{ij}(1 - \hat{p}_{ij})}{n_{ij}}} \, .$$

Since $\hat{n}_{oj}$ is a function of $\{\hat{p}_{ij}, \; i = 1,\ldots,k\}$ which are continuous at the points $\{p_{ij}, \; i = 1,\ldots,k\}$, and $\hat{p}_{ij} \xrightarrow{p} p_{ij}$, it follows that $\hat{n}_{oj}/n_{oj} \xrightarrow{p} 1$. Then it follows from $\hat{p}_j \xrightarrow{p} p_j$ that $\hat{\omega}_j/\omega_j \xrightarrow{p} 1$. When ω_j is replaced by $\hat{\omega}_j$ in (4.4), we have the following theorem for the set of three X^2 tests.

THEOREM 1 *The X^2 test statistics for $H_1(P_3; \; P_1)$, $H_2(P_2; \; P_1)$ and $H_3(P_3; \; P_2)$:*

$$X^2_{total} = \sum_j \frac{\hat{n}_{oj}(\hat{p}_j - \hat{p})^2}{\hat{p}_j(1 - \hat{p}_j)},$$

$$X^2_{linear} = \sum_j \frac{\hat{n}_{oj}(\hat{p}_j - \tilde{p}_j)^2}{\hat{p}_j(1 - \hat{p}_j)} \qquad and$$

$$X^2_{slope} = \frac{(\sum_j \hat{n}_{oj}(x_j - \bar{x}) \, / \, (1 - \hat{p}_j))^2}{\sum_j \hat{n}_{oj}(x_j - \bar{x})^2 \, / \, \hat{p}_j(1 - \hat{p}_j)}$$

are asymptotically distributed according to X^2_{N-1} if H_1 is true, X^2_{N-2} if H_2 is true and X^2_1 if H_3 is true, respectively. The three test statistics satisfy the identity: $X^2_{total} = X^2_{linear} + X^2_{slope}$.

If we estimate ω_j by $\ddot{\omega}_j = \hat{n}_{oj} / \hat{p}_o(1 - \hat{p}_o)$, where $\hat{p}_o = \sum_j \hat{n}_{oj}\hat{p}_j / \sum_j \hat{n}_{oj}$,

and replace ω_j of (4.4) by $\ddot{\omega}_j$, we have X^2 test statistics which are

similar to those of the theorem 1. Since $\ddot{\omega}_j/\omega_j \overset{P}{\longrightarrow} 1$ if H_1 is true,

we have the following theorem.

THEOREM 2 *The X^2 test statistics for $H_1(P_3; P_1)$, $H_2(P_2; P_1)$ and $H_3(P_3; P_2)$:*

$$X^{2\prime}_{total} = \sum_j \frac{\hat{n}_{oj}(\hat{p}_j - \hat{p}_o)^2}{\hat{p}_o(1 - \hat{p}_o)},$$

$$X^{2\prime}_{linear} = \sum_j \frac{\hat{n}_{oj}(\hat{p}_j - \tilde{p}_j)^2}{\hat{p}_o(1 - \hat{p}_o)} \quad and$$

$$X^{2\prime}_{slope} = \frac{(\sum_j \hat{n}_{oj}\hat{p}_j(x_j - \bar{x}))^2}{\hat{p}_o(1 - \hat{p}_o)(\sum_j \hat{n}_{oj}(x_j - \bar{x})^2)}$$

are asymptotically distributed according to X^2_{N-1}, X^2_{N-2} and X^2_1, respectively

if H_1 is true, and have the identity: $X^{2\prime}_{total} = X^{2\prime}_{linear} + X^{2\prime}_{slope}$.

We remark that $X^{2\prime}_{total}$, $X^{2\prime}_{linear}$ and $X^{2\prime}_{slope}$ are the adjustments of the

X^2 test statistics of Cochran and Armitage for stratification. But

$X^{2\prime}_{linear}$ is not asymptotically distributed according to X^2_{N-2} if H_2 is

true. If the denominator of X^2_{linear} is changed by $\tilde{p}_j(1 - \tilde{p}_j)$, we have

a X^2 test statistic which is similar to X^2_{linear}:

$$X^{2\prime\prime}_{linear} = \sum_j \frac{\hat{n}_{oj}(\hat{p}_j - \tilde{p}_j)^2}{\tilde{p}_j(1 - \tilde{p}_j)}.$$

Since $\tilde{p}_j \xrightarrow{p} p_j$ if H_2 is true, $X^{2\prime\prime}_{linear}$ is asymptotically distributed as X^2_{N-2} if H_2 is true. X^2_{total} and $X^{2\prime}_{total}$, X^2_{linear} and $X^{2\prime\prime}_{linear}$, X^2_{slope} and $X^{2\prime}_{slope}$ are asymptotically equivalent, respectively, so that we may also use a set of three statistics, $X^{2\prime}_{total}$, $X^{2\prime\prime}_{linear}$ and $X^{2\prime}_{slope}$, although they have no identity, that is, $X^{2\prime}_{total} \neq X^{2\prime\prime}_{linear} + X^{2\prime}_{slope}$.

4.1.2 Powers

In this section, by using the limiting powers under the appropriate sequences of alternative hypotheses, we give the approximate powers of our three tests.

THEOREM 3 *Against the null hypothesis of H_1, consider the following alternative:*

$$A_1; \quad p_j = \alpha + d_{1j} \quad (j = 1,\ldots, N),$$

where d_{1j} is nearly equal to zero. Then an approximation to the power of the X^2_{total}-test of level ε for A_1 is

(4.5) $Pr\{X^2_{total} > X^2_{N-1,\varepsilon} | A_1\} = Pr\{X^2_{N-1}(\lambda_t n_o.) > X^2_{N-1,\varepsilon}\},$

where

$$\lambda_t = \frac{1}{\alpha(1-\alpha)} \left[\sum_j d_{1j}^2 Q_{oj} - \left(\sum_j d_{1j} Q_{oj} \right)^2 \right],$$

$X_{r,\alpha}^2$; upper 100α % point of the X^2 distribution with r degrees

      of freedom,

$X_r^2(\lambda)$; non-central X^2 random variable with r degrees of freedom

      and non-centrality parameter λ,

$Q_{oj} = n_{oj}/n_{o\cdot\cdot}$.

PROOF Against the null hypothesis of H_1, consider the sequence of alternative hypotheses:

$$A_{n_o\cdot}; \quad p_j = \alpha + c_{1j}/\sqrt{n_o\cdot} \qquad (j = 1, \ldots, N).$$

Under $\{A_{n_o\cdot}\}$, the statistic $X_{total}^{2\prime}$, in the limit as $n_o\cdot$ tends to infinity subject to Q_{oj} remaining fixed, is distributed as the non-central X^2 distribution with $N - 1$ degrees of freedom and non-centrality parameter

$$\lambda_t' = \frac{1}{\alpha(1-\alpha)} \left[\sum_j c_{1j}^2 Q_{oj} - \left(\sum_j c_{1j} Q_{oj} \right)^2 \right].$$

Its proof is clear from Chapman and Nam [9]. Since $X_{total}^{2\prime}$ and X_{total}^2 are asymptotically equivalent under $\{A_{n_o\cdot}\}$, X_{total}^2's limiting power of level ε for $\{A_{n_o\cdot}\}$ is

(4.6) $\displaystyle \lim_{n_o\cdot \to \infty} Pr\{X_{total}^2 > X_{N-1,\varepsilon}^2 | A_{n_o\cdot}\} = Pr\{X_{N-1}^2(\lambda_t') > X_{N-1,\varepsilon}^2\}.$

Then, by approximating its power for A_1 by using (4.6), we have (4.5).

Q.E.D.

From the parallel discussions to the theorem 3, we have the following theorems.

THEOREM 4 *Against the null hypothesis of* H_2*, consider the following alternative:*

$$A_2; \quad p_j = \alpha + \beta x_j + d_{2j} \quad (j = 1, \ldots, N),$$

where d_{2j} *is nearly equal to zero. Then an approximate power of the* X^2_{linear}*-test of level* ε *for* A_2 *is*

$$(4.7) \qquad Pr\{X^2_{linear} > X^2_{N-2,\varepsilon} | A_2\} = Pr\{X^2_{N-2}(\lambda_\ell n_o.) > X^2_{N-2,\varepsilon}\},$$

where

$$\lambda_\ell = \sum_j d_{2j} z_j Q_{oj} - (\sum_j d_{2j} z_j)^2 / \sum_j z_j - (\sum_j d_{2j} z_j x_{oj})^2 / \sum_j z_j x^2_{oj},$$

$$z_j = \frac{Q_{oj}}{(\alpha + \beta x_j)(1 - (\alpha + \beta x_j))},$$

$$x_{oj} = x_j - \bar{x}.$$

THEOREM 5 *Against the null hypothesis of* H_3*, consider the following alternative:*

$$A_3; \quad \beta = d_3,$$

where d_3 *is nearly equal to zero. Then an approximate power of the*

X^2_{slope}-*test of level* ε *for* A_3 *is*

(4.8) $\qquad Pr\{X^2_{slope} > X^2_{1,\varepsilon} | A_3\} = Pr\{X^2_1(\lambda_s n_o.) > X^2_{1,\varepsilon}\},$

where $\qquad \lambda_s = d_3^2 [\sum_j Y_j x^2_{oj}] \quad$ *and* $\quad Y_j = \dfrac{Q_{oj}}{\alpha(1-\alpha)}.$

The above three theorems show that the three approximate powers are monotone increasing functions of n_o. because the three non-centrality parameters are linear functions of n_o.. Thus the larger the sample sizes, the larger the powers.

4.2 COMPARISONS BETWEEN THE PROPOSED METHOD AND OTHER STATISTICAL METHODS

In the previous sections, we discussed X^2 tests for a linear regression of the adjusted prevalence rates on the study factor intensities. On the other hand, when we are interested in a linear regression of the prevalence rates on the study factor intensities in each stratum, we can construct other X^2 tests based on the regression in each stratum.

In this case, three hypotheses parallel to $H_1(P_3; P_1)$, $H_2(P_2; P_1)$ and $H_3(P_3; P_2)$ are as fllows.

$K_1(P_3; P_1): \quad p_i \in P_3$ against $p_i \in P_1 - P_3$ for all i's,

$K_2(P_2; P_1): \quad p_i \in P_2$ against $p_i \in P_1 - P_2$ for all i's,

$K_3(P_3; P_2): \quad p_i \in P_3$ against $p_i \in P_2 - P_3$ for all i's,

where $p_i = (p_{i1}, \ldots, p_{iN})'$ and $p_i \neq p_{i'}$, $(i \neq i')$ in general.

4.2.1 Mantel's X^2-test

Mantel [5] proposed the X^2 test statistic given by

$$X_M^2 = \frac{(\sum_i \sum_j x_j (m_{ij} - e_{ij}))^2}{\sum_i \frac{n_{i\cdot} - m_{i\cdot}}{n_{i\cdot} - 1} (\sum_j e_{ij} x_j^2 - \frac{1}{m_{i\cdot}} (\sum_j e_{ij} x_j)^2)},$$

where $e_{ij} = m_{i\cdot} n_{ij} / n_{i\cdot}$, for the hypothesis of equi-prevalence rates against the ordered alternative for each stratum:

$$K_4(P_3; P_4): p_i \in P_3 \text{ against } p_i \in P_4 - P_3 \text{ for all } i\text{'s},$$

where $P_4 = \{p; p_1 \leq \ldots \leq p_N\}$ and $p_i \neq p_{i'}$, $(i \neq i')$ in general. This is also the test statistic for $K_3(P_3; P_2)$.

We shall compare X_{slope}^2 with X_M^2. We give first an approximation to the power of X_M^2 which can be proved by a discussion parallel to that in Section 4.1.2.

THEOREM 6 *Against the null hypothesis of K_3, consider the following alternative:*

$$A_4; \quad \beta_i = d_{4i} \quad (i = 1, \ldots, k),$$

where d_{4i} is nearly equal to zero. Then an approximate power of the

X^2_M-test of level ε for A_4 is

(4.9) $\qquad Pr\{X^2_M > X^2_{1,\varepsilon} | A_4\} = Pr\{X^2_1(\lambda_M n..) > X^2_{1,\varepsilon}\},$

where

$$\lambda_M = \frac{(\sum_i d_{4i} \sum_j Q_{ij} x^2_{ij})^2}{\sum_i \alpha_i (1 - \alpha_i)(\sum_j Q_{ij} x^2_{ij})},$$

$$Q_{ij} = n_{ij}/n..,$$

$$x_{ij} = x_j - \overline{x}_i.$$

Put an assumption that $p_i = p$ for all i's. Under the assumption, $H_3(P_3; P_2)$ includes $K_3(P_3; P_2)$. Thus we can compare the power of the X^2_{slope}-test with the power of the X^2_M-teat by comparing $\lambda_s n_o$. with $\lambda_M n...$ We define the ratio of $\lambda_s n_o$. to $\lambda_M n..$ by Λ. Since Λ is a complicated function of $\{Q_{ij}\}$, $\{w_i\}$, $\{d_{4i}\}$ and $\{\alpha_i\}$, a general comparison is difficult. Hence, we shall compare two tests in a simple situation.

THEOREM 7 *Put an assumption that* $p_i = p$ *for all i's. Against the null hypothesis of* K_3, *consider the following alternative:*

$$A'_4; \quad \beta_i = d_3 \quad (i = 1,\ldots, k).$$

Use the rates of the strata sizes to the total size as the standard population rates, that is, $w_i = n_i./n..$ $(i = 1,\ldots, k)$. *Then if*

(4.10) $\qquad n_{ij} = n_i. n._j/n..,$

it holds that $\Lambda = 1$ under $\dot{A_4}$.

PROOF Under the assumptions, $Q_{oj} = 1 \Big/ \sum\limits_{i} \dfrac{Q_i^2 \cdot}{Q_{ij}}$ holds. Thus Λ is a function of $\{Q_{ij}\}$:

$$\Lambda(\{Q_{ij}\}) = \sum\limits_{j} Q_{oj} x_{oj}^2 \Big/ \sum\limits_{ij} Q_{ij} x_{ij}^2.$$

The appendix shows that $Q_{oj} \leq Q_{\cdot j}$, with equality when $Q_{ij}/Q_i \cdot = Q_{i'j}/Q_{i'} \cdot$ ($i \neq i'$) which is equivalent to (4.10). Since (4.10) gives the equality:

$x_{oj} = x_{ij}$ ($i = 1, \ldots, k$), it holds that $\Lambda = 1$. Q.E.D.

4.2.2 Summing up X^2 Statistics in Each Stratum in the Style of Wood

Wood [11] summed up X^2 statistics in each stratum for the measure of fitness of linear models in each stratum.

By using "summing up X^2 statistics in each stratum" procedure, we can give three statistics for $K_1(P_3; P_1)$, $K_2(P_2; P_1)$ and $K_3(P_3; P_2)$:

$$X_1^2 = \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{N} \dfrac{n_{ij}(\hat{p}_{ij} - \hat{p}_i)^2}{\hat{p}_{ij}(1 - \hat{p}_{ij})} ,$$

$$X_2^2 = \sum\limits_{ij} \dfrac{n_{ij}(\hat{p}_{ij} - \tilde{p}_{ij})^2}{\hat{p}_{ij}(1 - \hat{p}_{ij})},$$

$$X_3^2 = \sum_i \frac{(\sum_j n_{ij}(x_j - x_i) / (1 - \hat{p}_{ij}))^2}{\sum_j n_i.(x_j - \bar{x}_i) / \hat{p}_{ij}(1 - \hat{p}_{ij})},$$

where
$$\begin{cases} \hat{p}_i = m_i./n_i., \quad \tilde{p}_{ij} = a_i + b_i x_j, \quad a_i = \hat{p}_i - b_i \bar{x}_i, \\ \bar{x}_i = \sum_j n_{ij} x_j/n_i. \text{ and } b_i = \sum_j m_{ij}(x_j - \bar{x}_i) / \sum_j n_{ij}(x_j - \bar{x}_i)^2. \end{cases}$$

The set of three statistics, X_1^2, X_2^2 and X_3^2 which will be called the individual X^2's, has the property that $X_1^2 = X_2^2 + X_3^2$, but contains no estimators of the adjusted prevalence rates. A parallel discussion as in Section 4.1.2 illustrates that the individual X^2's are approximately distributed according to the non-central X^2 distributions under the alternatives.

Put the assumption in 4.2.1, that is, $p_i = p$ for all i's. Under the assumption, $H_1(P_3; P_1)$, $H_2(P_2; P_1)$ and $H_3(P_3; P_2)$ include $K_1(P_3; P_1)$, $K_2(P_2; P_1)$ and $K_3(P_3; P_2)$, respectively. Put one more assumption that all w_i's are equal and $n_{ij} = n_{i'j}$ ($i \neq i'$, $j = 1, \ldots, N$). Under the adequate alternatives parallel to A_4', their non-central parameters are equal, respectively, to those of X_{total}^2, X_{linear}^2 and X_{slope}^2 which will be called the overall X^2's. Since the inequality:

$$Pr\{X_t^2(\lambda) \geq X_{t,\alpha}^2\} < Pr\{X_r^2(\lambda) \geq X_{r,\alpha}^2\}, \quad \text{if } t > r,$$

holds generally (c.f. Gupta and Perlman [10]), the powers of the individual X^2's are smaller than those of the overall X^2's because the former's

degrees of freedom are k times larger than the latter's.

We also expect that these approximations of the individual X^2's are less accurate than those of the overall X^2's respectively because the sample size of each stratum is smaller than the total sample size.

4.3 AN APPLICATION

Tsubota [1], [2] studied the relationship between air pollution and the prevalence rates of persistent cough and phlegm in Okayama Prefecture. We shall apply the procedure just discussed to his data. The survey districts consist of 12 communities. In each district, about 400 observations were collected from men and women in the range of 40 to 60 years of age. We employ the average concentration of NO_2 as an index of air pollution. For details, the original article may be referred to.

In this case, the confounding factors are sex, age and smoking habit. The age factor has two levels (40-49, 50-59). The factor of smoking has five levels (no smoking, ex-smoking, 1-10 cigarettes/day, 11-20 cigarettes/day and more than 20 cigarettes/day). Since female smokers are few in number, we combine ex-smokers and smokers for them. Fourteen strata are given in Table 4.1. The numbers of observations and cases who complain of persistent cough and phlegm are given in Table 4.2 and Table 4.3, respectively. The average concentrations of NO_2 are given in Table 4.4.

Table 4.1

14 strata of observations.

Number of stratum	Sex	Age	Smoking habit
1	male	40–49	no smoking
2	male	40–49	ex-smoking
3	male	40–49	1–10 cigarettes/day
4	male	40–49	11–20 cigarettes/day
5	male	40–49	more than 20 cigarettes/day
6	male	50–59	no smoking
7	male	50–59	ex-smoking
8	male	50–59	1–10 cigarettes/day
9	male	50–59	11–20 cigarettes/day
10	male	50–59	more than 20 cigarettes/day
11	female	40–49	no smoking
12	female	40–49	ex-smoking or smoking
13	female	50–59	no smoking
14	female	50–59	ex-smoking or smoking

Table 4.2

The number of observations.

| Stratum | \multicolumn{12}{c}{District} | Total | $n_i./n.. $ (%) |

Stratum	1	2	3	4	5	6	7	8	9	10	11	12	Total	$n_i./n..$ (%)
1	5	9	9	6	5	7	20	11	7	13	8	8	108	2.56
2	10	13	6	10	3	6	13	12	6	17	14	8	118	2.80
3	11	11	9	5	5	4	9	15	6	19	10	10	114	2.70
4	46	29	54	22	17	14	36	47	39	61	40	36	441	10.46
5	30	14	18	24	14	9	25	29	31	27	42	25	288	6.83
6	6	3	6	9	6	2	6	8	3	7	6	8	70	1.66
7	8	4	7	7	6	10	9	5	9	7	12	6	90	2.13
8	3	7	3	6	7	4	16	7	8	9	4	3	77	1.83
9	17	19	19	18	23	17	34	30	32	39	25	31	304	7.21
10	20	11	11	16	15	8	18	16	14	16	13	16	174	4.13
11	133	117	101	131	37	42	109	115	110	106	132	110	1243	29.47
12	11	5	10	11	5	7	13	8	5	18	22	5	120	2.84
13	91	81	77	101	56	39	77	79	81	83	76	102	943	22.36
14	13	12	4	9	14	4	17	16	7	13	15	4	128	3.03
Total	404	335	334	375	213	173	402	398	358	435	419	372	4218	100.00
Adjusted sample size n_{oj}	471.	317.	290.	369.	228.	197.	415.	425.	392.	465.	420.	407.	4395.	

Table 4.3

The number of cases who complain persistent cough and phlegm.

Stratum	District												Total
	1	2	3	4	5	6	7	8	9	10	11	12	
1	0	2	0	0	0	0	1	1	0	0	0	0	4
2	0	2	0	0	0	0	1	0	0	1	0	0	4
3	0	2	1	0	1	0	2	2	1	1	0	1	11
4	5	4	4	1	2	2	3	4	7	8	3	5	48
5	6	2	3	2	4	3	4	6	6	3	2	2	43
6	0	1	1	0	1	0	1	0	0	0	0	0	4
7	1	1	0	1	0	0	0	0	0	1	1	0	5
8	0	2	1	1	0	1	3	1	2	1	0	0	12
9	0	5	3	2	4	0	4	3	2	3	1	4	31
10	2	4	2	1	2	1	3	0	1	4	3	3	26
11	3	2	3	5	1	0	4	2	2	2	4	2	30
12	0	0	1	3	1	3	1	1	0	2	2	0	14
13	2	3	4	8	6	1	3	4	2	4	2	0	39
14	2	1	1	0	1	0	2	1	1	0	0	0	9
Total	21	31	24	24	23	11	32	25	24	30	18	17	280

Table 4.4

Average concentrations of NO_2 and estimated values of three types

of the prevalence rates in 12 districts.

	District											
	1	2	3	4	5	6	7	8	9	10	11	12
NO_2 (ppm)	.022	.030	.023	.023	.027	.022	.022	.020	.016	.020	.017	.016
Crude prevalence rate (%)	5.2	9.3	7.2	6.4	10.8	6.4	8.0	6.3	6.7	6.9	4.3	4.6
Adjusted prevalence rate (%)	4.8	10.5	8.1	6.4	9.8	6.5	7.6	6.1	6.4	6.5	4.3	4.5
Estimated prevalence rate (%)	6.7	9.6	7.1	7.1	8.5	6.7	6.7	6.0	4.6	6.0	4.9	4.6

We use the proportions of the stratum sizes to the total size as
the standard population rates. From these data, we have

$$p(x) = -0.01 + 3.56x,$$

$$\chi^2_{total} = 20.68^* \quad (11 \text{ d.f.}),$$

$$\chi^2_{linear} = 8.37 \quad (10 \text{ d.f.}),$$

$$\chi^2_{slope} = 12.31^{**} \quad (1 \text{ d.f.}).$$
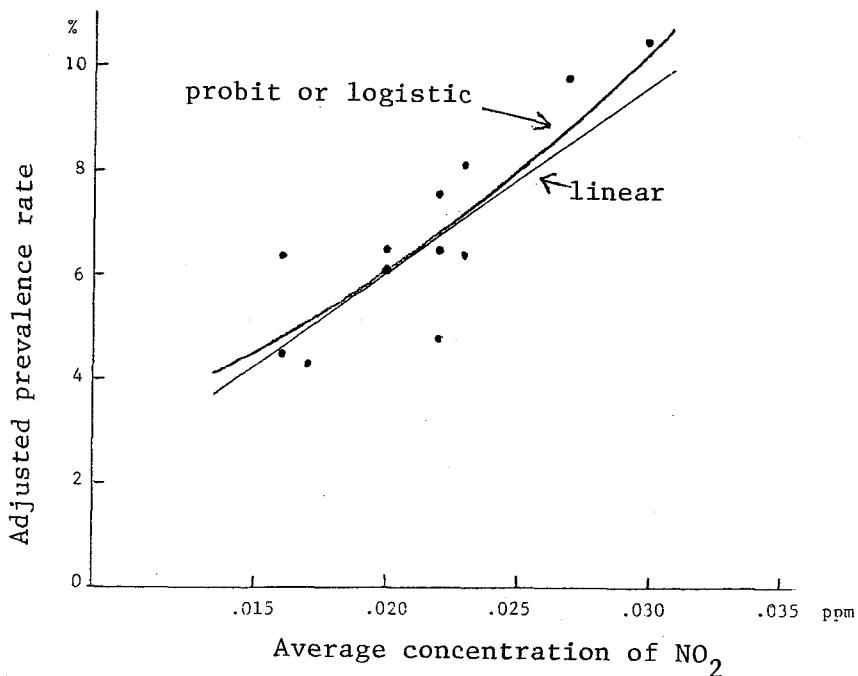
Fig. 4.1 The dose-response relationship between the average
concentration of NO_2 and the adjusted prevalence
rate of persistent cough and phlegm.

The results show that the difference of adjusted prevalence rates

at 12 districts are statistically significant ($p < 0.05$). It seems that

NO_2 index of air pollution explains this difference, for the linear model

is admitted and the regression coefficient is statistically significant

($p < 0.01$). The dose-response relationship is graphed in Fig. 4.1. To

compare X^2_{slope} with X^2_M, we also compute X^2_M and Λ:

$$X^2_M = 14.24**$$

$$\Lambda = 0.9327.$$

Thus, we can say that the difference between X^2_{slope} and X^2_M is negligible

in this case.

4.4 DISCUSSION

The situation where the prevalence rates of a disease may be affected

by a number of confounding factors is considered. The population in

each district is stratified according to the confounding factors. The

dose-response relationship between the study factor and the disease is

given by the linear regression model for the adjusted prevalence rates,

adjusted by the population composition, on the study factor intensities.

A set of three X^2 tests is proposed for an estimation and test procedure

of the linear regression. The main feature of this procedure is the

use of the adjusted prevalence rates and the adjusted sample sizes.

After computing these values, this procedure is similar to the procedure

of single stratum X^2 tests (Cochran [3] and Armitage [4]).

The X^2_M-test leads to no regression line and is useful only to test

the slope in the linear regression. Contrary to it, our procedure has

the following superiority.

1. It gives the regression line of the adjusted prevalence rates

    on the study factor intensities.

2. It gives a set of three X^2 tests for the regression line.

Compared with the individual X^2-tests, the overall X^2-tests of this

procedure have the following merits.

3. They are usually more powerful.

4. Their X^2 approximation may be more accurate.

As their subspecies, we have X^2_{total}, X^2_{linear}, X^2_{linear}'' and X^2_{slope}.

In the above discussion, we have assumed a linear regression line

for dose-response relationship. We can extend it to a general regression

curve defined by

$$(4.11) \qquad p(x) = \int_{-\infty}^{\alpha + \beta x} \phi(t) \, dt.$$

For example, $\phi(x) = \dfrac{1}{\sqrt{2\pi}} exp(- x^2/2)$ leads to a probit curve, and

$\phi(x) = exp(- x) \, / \, (1 + exp(- x))^2$ to a logistic curve. Here, we remark

that after a suitable transformation of the adjusted prevalence rates,

for example, probit transformation or logit transformation, the model

of (4.11) reduces to the linear regression model. Then, we can make

a parallel discussion as mentioned above. The effect of the transformation

appears only in the weight at each regression point. That is, the weights in Section 4.1.1 are changed into

$$(4.12) \qquad \omega_j = \frac{n_{oj}}{p_j(1 - p_j)\phi^2(p_j)}.$$

We shall apply the three X^2 tests under probit model or logistic model to the data in Section 4.3. The results are tabled in Table 4.5. Estimated probit curve and logistic curve are shown in Fig. 4.1 and they are nearly equal to each other in the data range.

Table 4.5 Estimated curve and three X^2 tests under probit model or logistic model.

Regression model	Estimated curve	X^2_{total}	X^2_{linear}	X^2_{slope}
probit	$p(x) = \int_{-\infty}^{-2.10 + 28.0x} \frac{1}{\sqrt{2\pi}} exp(-\frac{t^2}{2})dt$	21.96*	7.24	14.72**
logistic	$p(x) = \int_{-\infty}^{-3.88 + 57.4x} \frac{exp(-t)}{(1 + exp(-t))^2} dt$	22.05*	7.04	15.01**

* statistically significant at 5 %

** statistically significant at 1 %

APPENDIX

$$Q_{\cdot j}(\sum_i \frac{Q_{i\cdot}^2}{Q_{ij}}) - 1 = Q_{\cdot j}(\sum_i \frac{Q_{i\cdot}^2}{Q_{ij}}) - (\sum_i Q_{i\cdot})^2$$

$$= \sum_i Q_{i\cdot}^2 + \sum_i \frac{Q_{1j} + \cdots + Q_{i-1,j} + Q_{i+1,j} + \cdots + Q_{kj}}{Q_{ij}} Q_{i\cdot}^2 - \sum_i Q_{i\cdot}^2 - \sum_{i>i'}\sum 2Q_{i\cdot} \cdot Q_{i'\cdot}$$

$$= \sum_{i>i'}\sum (\sqrt{\frac{Q_{i'j}}{Q_{ij}}} Q_{i\cdot} - \sqrt{\frac{Q_{ij}}{Q_{i'j}}} Q_{i'\cdot})^2 \geqq 0,$$

then it holds that $1 \leqq Q_{\cdot j}(\sum_i Q_{i\cdot}^2/Q_{ij})$, that is, $Q_{oj} \leqq Q_{\cdot j}$. The equality sign holds when $Q_{ij}/Q_{i\cdot} = Q_{i'j}/Q_{i'\cdot}$ $(i \neq i')$.

REFERENCES

[1]　坪田信孝ら（1979）。　岡山県における呼吸器症状に関する疫学的研究、
日本公衆衛生学会誌、第26巻、第7号、333-350。

[2]　坪田信孝（1979）。　大気汚染と持続性せきたん有症率の関係、
日本公衆衛生学会誌、第26巻、第9号、481-492。

[3]　Cochran, W. G.(1954).　Some methods of strengthening the common X^2 tests, *Biometrics,* Vol. 10, pp. 417-451.

[4]　Armitage, P. (1955).　Tests for linear trends in proportions and frequencies, *Biometrics,* Vol. 11, pp. 375-385.

[5] Mantel, N. (1963). Chi-square tests with one degree of freedom; extensions of the Mantel-Haenszel procedure, *J. Amer. Statist. Ass.*, Vol. 58, pp. 690-700.

[6] Birch, M. W. (1964). The detection of partial association, *J. R. Statist. Soc. (B)*, Vol. 27, pp. 313-324.

[7] Birch, M. W. (1965). The detection of partial association, *J. R. Statist. Soc. (B)*, Vol. 27, pp. 111-124.

[8] Armitage, P. (1966). The chi-square test for heterogeneity of proportions, after adjustment for stratification, *J. R. Statist. Soc. (B)*, Vol. 28, pp. 150-163.

[9] Chapman, D. G. and Nam, J. (1968). Asymptotic power of chi-square tests for linear trends in proportions, *Biometrics*, Vol. 24, pp. 315-327.

[10] Gupta, S. D. and Perlman, M. D. (1974). Power of the noncentral F-test: Effect of additional variates on Hotelling's T^2-test, *J. Amer. Statist. Ass.*, Vol. 69, pp. 174-180.

[11] Wood, C. L. (1978). Comparison of linear trends in binomial proportions, *Biometrics*, Vol. 34, pp. 496-504.

ACKNOWLEDGEMENT